Department of Practice and Policy
UCL School of Pharmacy

PhD Thesis – University College London

Development of the Medicines Optimisation Assessment Tool (MOAT)

Targeting hospital pharmacists' input to reduce risks and improve patient outcomes

Cathy Anne Geeson

NIHR Clinical Doctoral Research Fellow
Pharmacy Department
Luton and Dunstable University Hospital

Declaration

I, Cathy Geeson confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Date:

Abstract

Background

Medicines optimisation is a key role for hospital pharmacists, but with ever-increasing demands on services, there is a need to increase efficiency while maintaining patient safety. The aim of this study was to use prognostic modelling to develop a prediction tool, the Medicines Optimisation Assessment Tool (MOAT^{TM}), to target patients most in need of pharmacists' input while in hospital.

Methods and analysis

Patients from adult medical wards at two UK hospitals were prospectively included into this cohort study between April and November 2016. Data on medication related problems (MRPs) were collected by pharmacists at the study sites as part of their routine daily clinical assessment of patients. Data on potential risk factors, such as polypharmacy and use of 'high-risk' medicines, were collected retrospectively from the information departments at the study sites, laboratory reporting systems and patient medical records. Multivariable logistic regression was used to determine the relationship between potential risk factors and the study outcome, namely preventable MRPs that were at least moderate in severity. A simplified electronic scoring system (the MOAT) was then developed.

Results

Among 1,503 eligible patient admissions, 610 (40.6%) experienced the study outcome. Eighteen risk factors were pre-selected for MOAT development, with 11 variables retained in the final model. The MOAT demonstrated fair predictive performance (concordance index 0.66), and good calibration. The decision threshold between 'low' and 'medium-risk' patients has a sensitivity of 90% (specificity 30%). The sensitivity for the threshold between 'medium' and 'high-risk' patients is 66% (specificity 61%). Decision curve analysis suggests that the MOAT has potential to be clinically useful across a wide range of predicted risk probabilities (from approximately 15% to 70%).

Conclusions

The MOAT has potential to predict those patients most at risk of moderate or severe preventable MRPs. External validation will be required to establish predictive accuracy in a new group of patients.

Impact statement

The work presented in this thesis has the potential to be put to beneficial use both outside and inside academia. Each of these is discussed below, together with proposed timescales and methods for impact realisation.

Impact on clinical practice

In terms of potential impact of this research outside academia, the main output is the Medicines Optimisation Assessment Tool (MOAT^{TM}) itself, which was developed with the aim of increasing the efficiency of hospital pharmacy services, reducing risks and improving patient outcomes. Once fully validated, the MOAT has the potential to impact on professional practice and patient safety.

The intention was for the MOAT to be adopted widely into clinical practice, therefore if generalisability and clinical effectiveness are demonstrated, the MOAT has the potential to be used across the UK, and potentially more widely; I have already received an expression of interest from a pharmacist in Australia regarding external validation of the MOAT within her clinical setting.

Impact within this research area

This study has also created knowledge that may inform future research in this field. This includes:

- consensus views on potential risk factors associated with medication related problems (MRPs) in hospitalised patients;
- data on MRP occurrence in adult medical patients in UK hospitals. More specifically, the prevalence of MRPs and moderate or severe preventable MRPs, and their breakdown by MRP sub-categories;
- quantification of the potential variability in MRP identification by hospital pharmacists;
- the views of practising pharmacy clinicians on the requirements of a predictive tool, including presentation and usability.

My academic supervisors at UCL have offered a place to a prospective PhD candidate to progress work in this area; if accepted, I have been asked to be a clinical advisor.

Realisation of impact

While the academic knowledge created by this research has potential to be used immediately, further research will be required prior to implementation of the MOAT into

routine practice. This will include external validation to assess predictive accuracy in a new sample of patients, and impact and implementation studies. I estimate this will take three to four years, dependant on funding opportunities.

Initially I intend to raise awareness of the findings of this study through dissemination via presentations at professional, academic and scientific meetings and conferences, and submission for publication in peer-reviewed journals. I also plan to present at relevant patient / public meetings at the study sites, and to work with the patient and public members of the project steering group to develop a wider public dissemination strategy. I hope to secure further funding to validate the MOAT, and once fully validated I would aim to work with key decision makers, such as the Head of Research at the Royal Pharmaceutical Society, and Medication Safety Team at NHS England to advise on further dissemination and adoption of the MOAT into practice. There is also potential to work with software developers to integrate the MOAT into automated systems such as electronic health records systems, permitting automated risk assessments in ‘real-time’, further supporting implementation into clinical environments.

Acknowledgments

First, I would like to give my heartfelt thanks to Professor Bryony Dean Franklin and Dr Li Wei for their unwavering support, guidance and encouragement. It has been an absolute privilege to work under your supervision. In particular, I would like to thank Professor Bryony Dean Franklin who is a truly inspirational role model!

I am sincerely grateful to Dr Mary Evans and Lindsay Smith for their clinical supervision, and for permitting this research to take place within their organisations. I would also like to thank the pharmacy staff involved in data collection; your dedication and combined enthusiasm were a tremendous support. Additional thanks go to Jack Glendenning, Colin Merrill, Andy Finch and Alan Osman for their support with technical aspects of data collection. I am also particularly grateful to Sehrush Hussain and Shahnaz Begum for their assistance with the simulated medication related problem (MRP) identification exercise, and to members of the expert panel, Sue Lee, Dr Siva Puthrasingam and Ann Williams; your expertise and experience were greatly appreciated. My sincere thanks also go to Aneesh Khurana for his technical input into the development of the electronic Medicines Optimisation Assessment Tool (MOAT^{TM}) scoring system. Similarly, I am indebted to the pharmacists and clinical pharmacy technicians who participated in the assessment of the MOAT's clinical credibility.

I would also like to recognise the patient and public members of the MOAT project steering group: Helen Clothier, Marie-France Capon, Brian Smith, Derek Wright and Tom Drabble. These amazingly supportive people ensured that a patient / carer perspective remained prominent during the research, while providing encouragement and valuable advice throughout.

I would also like to thank the National Institute for Health Research for the opportunity to undertake a Clinical Doctoral Research Fellowship. This enabled me to fully concentrate on my academic and clinical development, and afforded access to outstanding training and development opportunities, for which I feel truly fortunate.

Finally, to my friends and family, your unquestioning support and understanding has enabled me to maintain my passion, my vision and my focus. I am particularly grateful to my wonderful parents, who helped mould me into the person I have become, while always encouraging me to follow my dreams. I also want to thank my fantastic children, Toby and Ellie; you brighten my life more than you could know!

I owe and dedicate this thesis to you all.

Table of contents

List of tables

List of figures

Abbreviations

ADE	Adverse drug event
ADR	Adverse drug reaction
AIC	Akaike Information Criterion
AIDS	Acquired immunodeficiency syndrome
ALP	Alkaline phosphatase
ALT	Alanine aminotransferase
ART	Assessment of Risk Tool
AST	Aspartate aminotransferase
ATC	Anatomical Therapeutic Chemical
BD	'*Bis in die',* meaning 'twice daily'
BS model	Backward selection model
C-index	Concordance index
CD	Controlled drug
CDR	Clinical decision rule
CHAIN Network	Contact, Help, Advice and Information Network
CHARMS	CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies
CI	Confidence internal
CINAHL	Cumulative Index of Nursing and Allied Health Literature
CIRS	Cumulative illness rating scale
dd	Deviance statistic
df	Degrees of freedom
DRG	Diagnosis-related group

DRP	Drug related problems (also known as medication related problems / MRPs)
dx2	Chi-square fit statistic
EEPRU	Policy Research Unit in Economic Evaluation of Health and Care Interventions
EMBASE	Excerpta Medica dataBASE
EPV	Events per variable
FMI	Fraction of Missing Information
FP	Fractional polynomial
GGT	Gamma-glutamyl transpeptidase
HbA1c	Haemoglobin A1c / glycated haemoglobin test
HIV	Human immunodeficiency virus
HRA	Health Research Authority
ICE	Sunquest Integrated Clinical Environment
INR	International Normalised Ratio
IPA	International Pharmaceutical Abstracts
IQR	Interquartile range
ISMP	Institute for Safe Medication Practices
ITU	Intensive therapy unit
IV	Intravenous
LFTs	Liver function tests
Lowess	Locally weighted scatterplot smoothing
MAR	Missing at random (missingness mechanism)
MCAR	Missing completely at random (missingness mechanism)
MDC	Major diagnostic category

MDRD	Modification of Diet in Renal Disease
ME	Medication error
MEDLINE	Medical Literature Analysis and Retrieval System Online
MERIS	Medicines risk score
MFP	Multivariable fractional polynomial
MICE	Multiple imputation by chained equations
MMPT	Medicines management pharmacy technician
MNAR	Missing not at random (missingness mechanism)
MOAT	Medicines Optimisation Assessment Tool
MRCI	Medication Regimen Complexity Index
MRP	Medication related problem (also known as drug related problems / DRPs)
N/A	Not applicable
NGT	Nominal group technique
NHS	National Health Service
NIHR	National Institute for Health Research
NOAC	Novel oral anticoagulant
NSAID	Non-steroidal anti-inflammatory drug
PAL	Patient acuity level
PAST	Pharmaceutical assessment screening tool
PCNE	Pharmaceutical Care Network Europe
PR	Per rectum
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PRN	'*Pro re nata*' meaning 'as required'

PROGRESS	PROGnosis RESearch Strategy
PV	Per vagina
QDS	'*Quarter die sumendum*', meaning 'four times daily'
QUIPS tool	Quality in Prognostic Studies tool
REC	Research Ethics Committee
ROC	Receiver Operating Characteristic
RVI	Relative Increase in Variance
SD	Standard deviation
SR	Sustained release
STAT	'*Statum*', meaning 'to be given immediately'
STOPP	Screening Tool of Older People's potentially inappropriate Prescriptions
TDM	Therapeutic drug monitoring
TRIPOD	Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
TTA	Discharge prescription ("To Take Away")
UCL	University College London
UK	United Kingdom
USA	United States of America
UTI	Urinary tract infection
VIF	Variance inflation factor

Chapter 1: Overview

This thesis describes the development of a prediction tool, the Medicines Optimisation Assessment Tool (MOAT™), to identify patients at highest risk of moderate or severe preventable medication related problems (MRPs) during admission to a hospital medical ward. The thesis is presented in 11 chapters; each is briefly described below.

Chapter 1 – Overview

This comprises a brief summary of each subsequent chapter.

Chapter 2 – Background

The background provides the medical context and rationale for development of the MOAT. It comprises evidence supporting the need to reduce avoidable medication-related harm, the role played by hospital clinical pharmacists in medicines optimisation, the need for clinical prioritisation, and the potential use of 'prediction research' to permit effective targeting of patients.

Chapter 3 – Literature review

The literature review summarises existing evidence related to the prediction of adverse medication-related outcomes. This includes the identification of potential risk factors, known as prognostic factors, and a critical appraisal of existing prediction tools. The review resulted in the identification of 59 possible prognostic factors. It also suggested that the currently available prediction tools are not suitably robust for routine clinical use in terms of validated predictive accuracy, and/or generalisability.

Chapter 4 – Aim and objectives

The aim and objectives for the research were developed to address the gap in the evidence base identified in the preceding chapters. In summary, the aim was to develop an evidence-based prediction tool (using prognostic modelling), which has potential to increase the efficiency of hospital pharmacy services, reduce risks and improve patient outcomes.

Chapter 5 – Selection of candidate predictors

Chapter 5 describes the selection of potential prognostic factors (known as candidate predictors) to be used for MOAT development. This involved consideration of evidence from previous research, consensus of clinical experts and patient / public representatives, and a review of suitability in terms of the methodological requirements for prognostic modelling. One hundred and eighteen potential predictors were

considered, 59 from previous research and 59 suggested by clinical experts; 18 were pre-selected for use in MOAT development.

Chapter 6 – Data collection for model development

Chapter 6 describes the methodology, methods, and results of data collection. This includes the selection of study participants, flow of patients through the study, key characteristics of study admissions, analysis of missing data, data entry checks, descriptive analysis of MRP data, and ethical considerations.

Chapter 7 – Pharmacists' identification of MRPs: a validation exercise

A potential limitation of MRP data collection, identified in chapter 6, was the possible impact of knowledge, experience and skills of pharmacists on their ability to identify potential MRPs. Chapter 7 describes the assessment of this potential variability through use of a simulated MRP identification assessment exercise.

Chapter 8 – Analysis of outcome events

Chapter 8 describes the severity assessment of MRPs, which was undertaken to identify patients with the outcome event of interest, namely at least one moderate or severe preventable MRP. Chapter 8 also includes descriptive analyses of the outcome events identified.

Chapter 9 – Modelling

Chapter 9 describes the use of multivariable logistic regression modelling to develop the MOAT. This chapter includes sample size calculation, exploratory data analysis, use of multiple imputation, model development, assessment of model performance, adjustment for optimism (to reduce potential for overconfident predictions when applied to a new group of patients), and development of an electronic decision aid.

Chapter 10 – Assessment of the MOAT's clinical credibility

Assessment of the potential usability of the MOAT in clinical practice is described in chapter 10. This includes a review of the content validity, ease of use, potential workload implications, and potential clinical risk associated with false negative predictions.

Chapter 11 – Overall discussion

The overall discussion considers the key results of this research with reference to the study's objectives. Previous literature and limitations are also discussed. This is followed by implications for future research, and an overall conclusion.

Chapter 2: Background

Medication-related harm

Medicines play a crucial role in maintaining health and are the most common intervention in healthcare. However, in the United Kingdom (UK), as elsewhere, there is a growing body of evidence that there is a need to improve medicines use[1-7]. This includes the Francis and Berwick reports[1 2], which call for a number of actions to improve patient safety and reduce avoidable harm, and the World Health Organization's Global Patient Safety Challenge: Medication Without Harm[7], which was published in 2017, and outlines their global initiative to reduce the level of severe, avoidable medication-related harm by 50% over five years.

Various terms are used for adverse medication-related outcomes, including adverse drug reactions (ADRs), adverse drug events (ADEs), and medication errors (MEs). The relationship between each of these is shown in Figure 1[i] (produced by Otero *et al*[8]), which shows that ADRs are effectively a subset of ADEs, representing non-preventable medicines-related harm. MEs span all preventable events, and may or may not result in harm.

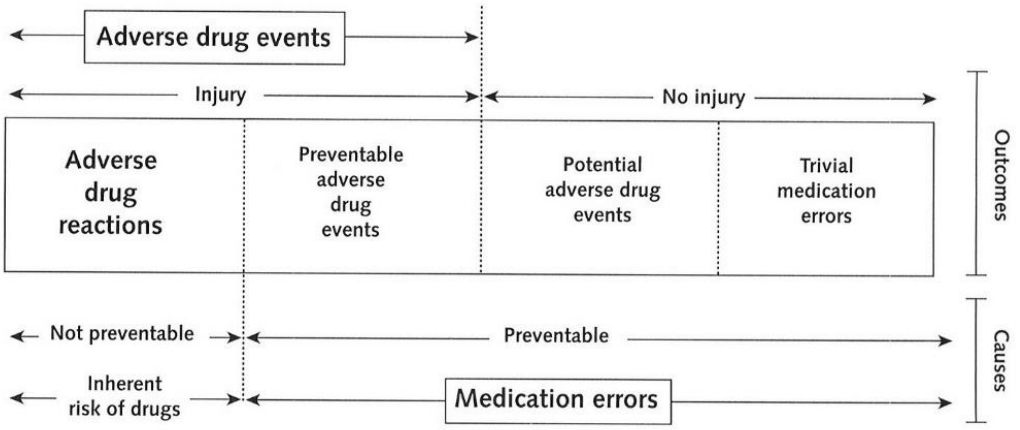


Figure 1 – Relationships between adverse drug events, adverse drug reactions and medication errors

Research in the UK has shown that 1 in 20 prescription items in general practice contain either a prescribing or monitoring error[9]; 5-8% of all hospital admissions are due to preventable adverse effects of medicines[10]; 30-70% of patients have an error or unintentional change to their medicines when care is transferred[3]; and there is a prescribing error rate in hospitals of almost nine errors per 100 medication orders[11].

[i] *Permission to use this image obtained from Annals of Internal Medicine*

This has human and financial costs[5 12-14]. Research has shown that admissions related to ADEs cost the National Health Service (NHS) up to £466m annually, with most being avoidable[10]. Focusing specifically on hospitalised patients, it is estimated that the annual cost to treat preventable ADEs in a 400-bed acute hospital in the UK is £0.3m-£1m[15], and a systematic review found that ADEs in hospitalised patients prolong hospitalisation by 3.4 days[16].

Medicines optimisation

Historically, ADEs have been the focus of studies on medication-related harm[12], but problems can also result from suboptimal medicines use, such as ineffective treatments or subtherapeutic doses. It is estimated that only 4-21% of patients in primary care receive optimum benefit from their medicines[17], and it has been suggested that research efforts should also identify patients with unrealised benefits[18]. A term that encompasses both aspects is medication related problems (MRPs), defined as all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome (from prescribing to administration)[12 16 19 20]. This also shifts the focus from 'medicine-related harm' to 'medicines optimisation', which can be described as the safe and effective use of medicines to enable the best possible outcomes[5]. Medicines optimisation is high on the English national agenda, with guidance issued by the Royal Pharmaceutical Society and the National Institute for Health and Care Excellence[4 5].

Hospital clinical pharmacists

Medicines optimisation is a key function of pharmacists[21-24], and a number of systematic reviews conclude that addition of clinical pharmacist services to the care of hospital inpatients generally improves the quality, safety, and efficiency of patient care[16 25 26]. Ideally, pharmacists would see every patient daily, but medicines optimisation is not the only goal for hospital pharmacy services in England[21 27]. Other service developments are required, such as the delivery of 7-day services[28] and the Hospital Pharmacy Transformation Programme, as set out in the recent review by Lord Carter on improving productivity and performance in English NHS acute hospitals[29]. Owing to the financial challenges that face the NHS, these developments often have to be achieved within existing funding through increased efficiency and innovation[5 30 31]. There have therefore been calls from international government organisations and professional bodies for effective ways for pharmacy services to target patients most in need[23 32-36].

Prioritisation / prediction tools

Clinical prioritisation has been proposed as a way to permit pharmacy services to focus on the greatest need and where clinical pharmacy input is likely to have greatest impact. This requires a method to triage patients to assign 'pharmaceutical acuity'[36 37]. There are recognised risk factors for MRPs, for example polypharmacy, renal impairment, and the use of 'high-risk' medicines[38], but to target patients appropriately pharmacists need to be able to apply this knowledge effectively and consistently within their routine clinical practice.

Predicting clinical risk is well established in medicine. Tools such as cardiac-risk calculators[39], and the Waterlow score[40] (shown in Appendix A2.1), which assesses the risk of pressure ulcers, are both used daily across the NHS. Prediction tools to identify hospitalised patients at risk of adverse medication-related outcomes have been developed[41-52], but the majority identify patients at risk of ADRs[41-43], ADEs[44 45], or MEs[47], rather than MRPs, or are based on 'expert opinion' rather than statistical determination[46 48-51].

Prediction research

Interest in prediction research (also known as prognosis research), has developed rapidly in recent years. It involves use of statistical methods to predict future health outcomes among people with a given baseline health status, and therefore has potential to inform clinical decision making, improve patient care, and make healthcare more efficient[53 54]. Prognostic modelling is one component of prognosis research, in which multiple risk (prognostic) factors are statistically combined to predict future clinical risk for an individual patient[55]. However, many published prognostic model studies have been criticised in terms of methodological shortcomings, limiting their reliability and applicability[55 56], as well as poor reporting, which limits the ability to effectively assess the risk of bias[54 57]. Both problems ultimately limit the usefulness of the prognostic models. The perceived inadequacies in prognostic model research prompted the recent publication of recommendations for prognosis research by the PROGnosis RESearch Strategy (PROGRESS) partnership[53 55 58 59], together with specific guidelines for reporting[54 57] and critically appraising[60] prognostic model research.

Gap in the evidence base

An initial review of the literature suggested that a methodologically sound prognostic model that can target hospital patients most in need of pharmacists' input, based on their risk of MRPs, does not yet exist. My aim was, therefore, to address this gap in the evidence base by developing a suitable model, the Medicines Optimisation Assessment Tool (MOAT^{TM}). This was driven by my desire to reduce risk, improve outcomes, and increase efficiency of hospital clinical pharmacy services, thereby supporting delivery of national targets related to patient safety, medicines optimisation, and service provision.

The next stage was to review the literature to: (1) review previous research into the development of prediction tools to identify hospitalised patients at risk of adverse medication-related outcomes; and (2) identify potential prognostic factors. This is described in the next chapter.

Chapter 3: Literature review

3.1 Introduction

As described in the previous chapter, the aim of this study was to develop a prediction tool, the Medicines Optimisation Assessment Tool (MOAT™), using prognosis research methods. Prognosis research can be subdivided into distinct, but interrelated themes[48]. These include the identification of factors associated with prognosis, known as prognostic factor research, and prognostic model research, where prognostic factors are combined to predict future clinical risk for individual patients[53 55].

A key purpose of this literature review was to critically appraise the existing prediction tools[41-52]. In 2014 Stevenson *et al*[61] published a systematic review of adverse drug reaction (ADR) and adverse drug event (ADE) prediction models for hospitalised older adults. My aim, therefore, was to update Stevenson's review, and broaden it to include prediction tools for other adverse medication-related outcomes, those developed using 'expert opinion', and those developed for adult patients irrespective of age, as these were excluded in Stevenson's review.

Prognostic factor research can guide model development by informing the selection of potential prognostic factors[62 63], therefore the second purpose was to inform the selection of potential prognostic factors for the MOAT (discussed further in chapter 5, section 5.2.1). After undertaking my initial database searches, Suggett *et al*[64] published a systematic review of risk factors associated with the need for pharmaceutical interventions in a hospital setting. Our research question and eligibility criteria were broadly similar, but our reviews differ in search close date (Suggett's search closed in July 2013). My review therefore provides an update, and an opportunity to compare and contrast findings.

In summary, the objectives were to review studies that:

- developed prediction tools for adverse medication-related outcomes for adult patients admitted to hospital medical wards, to appraise the quality, applicability and limitations of existing prediction tools;
- assessed the association between prognostic / risk factors and this outcome event, to identify potential prognostic factors to be used in MOAT development.

3.2 Methods

To minimise bias, I established the research question and inclusion criteria *a priori*[65] using PICOS criteria (participants, interventions, comparators, outcomes, and study design)[66]. Three elements of the PICOS criteria were relevant to this review (participants, outcomes, and study design). I also followed the guidance developed by Moons *et al*[60] for the critical appraisal and data extraction for systematic reviews of prediction modelling studies (shown in Appendix A3.1). This provides advice on framing the review question, search strategy, and eligibility criteria.

The format and content of the review are based on the Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA) guidelines[67].

3.2.1 Eligibility criteria

The eligibility criteria, with rationale, are summarised below:

Type of studies: Existing research in this field includes a wide range of study types, I therefore included observational studies (cohort, case-control, and cross sectional studies), and consensus studies (expert opinion, Delphi method, and nominal group technique). No types of study were excluded. I restricted the search to published studies, hence excluded grey literature (reports, theses, opinion pieces, and conference proceedings)[68].

Types of participants: To reflect the proposed target population for the MOAT I selected studies of adult inpatients (defined as 18 years and over) on general medical wards. Studies that only included the following specialised groups / treatments were excluded:

- patients on specialised units, including intensive care and cardiology;
- patients with specific conditions, including renal disease, human immunodeficiency virus (HIV) infection, diabetes, cancer, depression, and dementia;
- patients treated with a single group of medicines, such as antibiotics.

Studies that included specialised patients as part of a wider group, including those on general medical wards, were included in the review.

Type of outcome measures: I included studies that identified prognostic factors, and/or developed prediction tools, for patients at risk of adverse medication-related outcomes during hospital admission. Studies related to medication problems causing admission to hospital, or occurring following discharge, were excluded as they fall

outside the scope of the current research. Various outcomes measures have been used in these studies, including ADRs, ADEs, medication related problems (MRPs), and medication errors (MEs). A range of definitions are also used for each outcome, making direct comparison among studies difficult. I therefore included studies of all the above outcomes events.

Language/publication date: For practical reasons I restricted studies to those published in English. No publication date restriction was applied.

3.2.2 Information sources

As the identification and resolution of adverse medication-related outcomes is a multi-professional discipline I searched a range of databases: Excerpta Medica dataBASE (EMBASE); Medical Literature Analysis and Retrieval System Online (MEDLINE); Cumulative Index of Nursing and Allied Health Literature (CINAHL), and International Pharmaceutical Abstracts (IPA). Reference lists of relevant studies were also scanned and experts in the field consulted to check for other relevant studies.

The database searches were initially completed between August and November 2015, with weekly electronic update alerts checked to identify subsequent publications.

3.2.3 Search

To allow for variation in index terms, I searched each database separately. The research question was divided into facets, with synonyms used to account for the different terminology, and truncation symbols to retrieve variant spellings and word endings. The facets were then added together using the Boolean AND operator. Three facets were used, the first included the outcome measure (MRPs, ADRs, ADEs, and MEs), the second covered risk prediction, so included risk factor, predictor variable, prediction tool, clinical decision rule, prioritisation tool and prognostic model. The third was used to capture studies of hospitalised patients. The full electronic search strategy used for EMBASE is shown in Appendix A3.2.

3.2.4 Study selection

I screened the titles of all records identified by each database search and excluded records according to the eligibility criteria. Abstracts and full-text publications were reviewed for potentially relevant records. Duplicate records were removed, and duplicate publications identified (i.e. studies published more than once). All reports from duplicate publications were considered and only the paper related to the original

study to develop the prediction tool or identify prognostic factors selected for the review.

3.2.5 Data collection

I developed and piloted data extraction records for six sets of data. The extraction records and data items are listed below.

All studies:

1. **General data**. This included: research method (consensus or statistical); country where study undertaken; sample size; age of participants; patient group (i.e. speciality / specialities studied); whether study was carried out prospectively or retrospectively; type of outcome event (MRP, ADE, ADR, ME); whether the outcomes were rated for severity and preventability (for the purposes of selecting the outcomes for statistical analysis); whether a prediction tool was developed.
2. **Data on the risk / prognostic factors studied**. For consensus studies this included which factors were initially considered, and which were subsequently selected as being associated with the outcome event. For statistical studies it included details of all factors studied, and whether a univariable and/or multivariable association was found between the prognostic factor and outcome event.

Consensus studies:

3. A summary of study purpose (identification of risk factors or development of a prediction tool); study method (expert opinion, Delphi method, nominal group technique); details of the expert group; results of testing / validation.

Statistical studies:

4. **Prognostic factor studies**: data on the six domains of the Quality In Prognostic Studies (QUIPS) tool[69].
5. **Prognostic modelling studies**: data on the 11 domains of the 'CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies' (CHARMS)[60].

Prediction tool studies (consensus and statistical):

6. Method of development; type of tool developed; prognostic / risk factors included in the prediction tool; scoring system; results of testing / validation.

There was overlap between data required to assess the risk of bias within individual studies, and for the overall synthesis of results. For example, testing / validation of

prediction tools forms part of the risk of bias review for prognostic model and consensus studies individually, but also forms part of the overall synthesis of results. For ease of reference I reported these data in each relevant section of the results.

3.2.6 Risk of bias in individual studies

To assess the risk of methodological bias within individual studies I used the QUIPS tool[69] and CHARMS[60] checklist to assess the prognostic factor and modelling studies respectively. The consensus studies were assessed based on consideration of the appropriateness of the method, expert group, and results from testing / validation where available. The results were used when summarising the available evidence.

The QUIPS tool (Appendix A3.3) permits prognostic factor studies to be rated as high, moderate or low bias for each of six domains (study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis and reporting) using explicit criteria for each domain. As the purpose of this literature review was to identify potential prognostic factors to include in development of the MOAT rather than provide a definitive answer as to which factors are prognostic, I chose to use the QUIPS tool to provide a high-level overview of the included studies rather than a detailed critique.

The CHARMS checklist (Appendix A3.4) lists the key data items to extract for systematic reviews of prediction model studies to identify potential sources of bias, and issues that may affect the applicability of the model. The checklist is grouped into 11 domains (source of data, participants, outcome, candidate predictors, sample size, missing data, model development, model performance, model evaluation, results, interpretation and discussion). I chose to use the CHARMS checklist to complete a detailed review because of the importance of fully understanding the strengths and limitations of prognostic models currently available in this field.

3.2.7 Summary measures / synthesis of results

The analysis of individual studies was split into two sections, prognostic factor, and prediction tool development studies, as described below.

Prognostic factor studies

To inform the selection of prognostic factors for the MOAT I performed a simple count of the number of studies where an association was found between the factor and outcome event, a method recommended by Steyerberg[70]. Correlation between variables can affect the results of univariable analysis, either increasing or decreasing

the association[70], I therefore chose to use univariable and/or multivariable associations as the outcome event of primary interest. Statistical combination of data was not performed due to differences in study design and outcome event measures. Steyerberg also advises caution when combining data due to publication bias[70], in which authors only report results when associations are found, so biasing effects to more extreme values.

For each study I recorded the prognostic factors studied, and whether an association was found. I categorised some prognostic factors into groups to simplify the analysis, for example where individual diagnoses were studied I recorded this under the group 'diagnosis'. I then combined studies by study method (statistical and consensus), and outcome measure (MRP, ME, ADE, ADR) to establish the number of studies within each category that found an association. I also combined the results from all studies to give an overall indication of the level of association, in terms of the proportion of studies that found an association between the prognostic factor / group and the outcome event.

Prediction tool development studies

For each prediction tool development study, I recorded the method of development (consensus or statistical), type of tool developed (electronic scoring algorithm, or risk model / score), prognostic / risk factors included in the prediction tool, scoring system, and results of testing / validation.

3.2.8 Risk of bias across studies

Assessing bias across prognosis research studies is challenging as there is currently no requirement for study registration, making it difficult to assess publication or selective reporting bias[54]. Assessment of selective reporting bias is also hampered by deficiencies in the quality of reporting of prognosis research, with key details often missing, such as which prognostic factors were examined[57]. For these reasons the risk of bias across studies was not considered as part of this literature review.

3.3 Results

3.3.1 Study selection

Figure 2 shows the PRISMA flow diagram of study selection. The database searches identified 5,020 records. The initial numbers identified from each database were:

- EMBASE 2724 records;
- MEDLINE 1280 records;
- CINAHL 185 records;
- IPA 831 records.

An additional record was identified through a report in the pharmaceutical press[71], and six records from subsequent weekly electronic update alerts[43 46 47 51 72 73]. Titles / abstracts were screened and the potentially relevant records combined (total 387). Thirty nine duplicates records were removed, leaving 348 records. Each record was reviewed and a further 276 excluded. The majority of these reports (274) did not meet the eligibility criteria. Two reports were potentially eligible but were not retrievable (from the UCL Library Service or NHS Evidence)[74 75]. Full-text for 79 articles were retrieved and reviewed. Duplicate publications were identified and the paper related to the original study selected for the review. A further assessment against the eligibility criteria was then made.

A total of 31 studies were identified for inclusion in the literature review. Twelve of the studies involved the development of a prediction tool (in addition to identifying prognostic factors); the remaining 19 involved the identification of prognostic factors only.

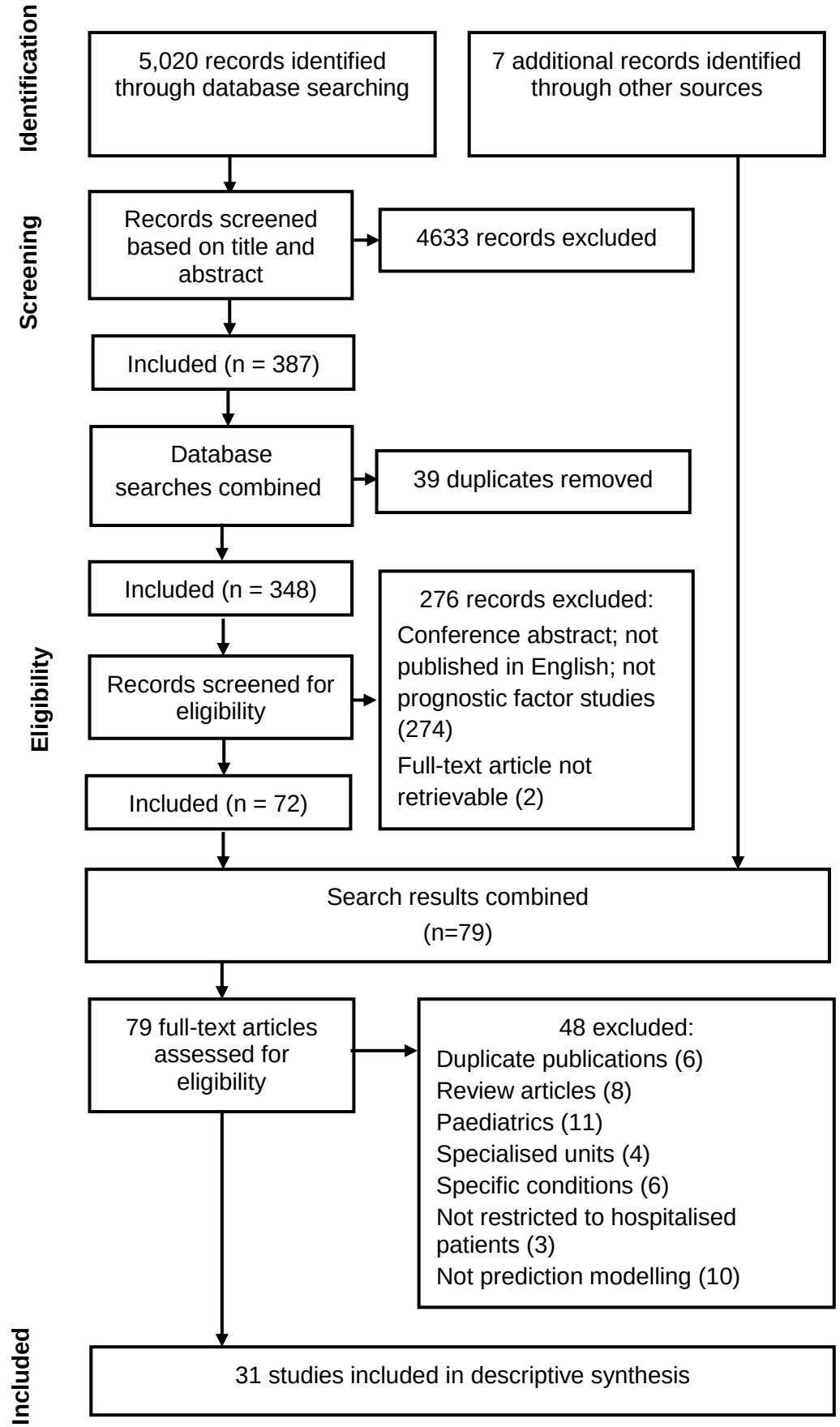


Figure 2 – PRISMA flow diagram showing the stages of the literature review

3.3.2 Study characteristics

The characteristics of studies selected for the review are shown in Appendix A3.5, and summarised below. All studies identified risk / prognostic factors, with a subset (12 studies) also developing prediction tools. I have therefore separated this section into a review of all studies, followed by a review of prediction tool studies only.

3.3.2.1 All studies

Type of studies: Six of the 31 studies used consensus methods. Of these, five developed prediction tools for medication-related harm, and the other identified risk factors only. The remaining 25 used statistical methods; seven were prognostic model studies, and 18 identified prognostic factors only (summarised in Table 1).

Of the 25 statistical studies, 19 were prospective and six retrospective. Twenty two used logistic regression. Log-linear, Cox, and Poisson regression were used in the remaining three.

Types of participants: All 31 studies included patients on hospital medical wards / specialities, but 18 also included patients from other specialities (surgery, orthopaedics, maternity, intensive care, cardiology, gynaecology, and urology). All studies included only adult patients, with six restricted to older adults (aged 65 years and over). The number of patients involved in each of the statistical studies ranged from 131 to 68,835, with a total of 250,585 patients included in the 25 studies.

The majority of the studies (22) involved research carried out in Europe (four in the Netherlands; four in England; two in Ireland, Spain, Switzerland, France and Germany; one in Scotland, Italy, Denmark and Norway). Five studies were conducted in North America (three in the USA and two in Canada); two in Africa (Ethiopia and Uganda); one in Asia (Israel), and one in Australasia (New Zealand).

The studies were published between 1997 and 2017.

Type of outcome measures: All studies involved the identification of risk / prognostic factors, and 12 also developed prediction tools. Nine studies used MRPs as the outcome measure, eight used ADEs, eight ADRs, five MEs, and the remaining study used ADEs and MEs (see Table 1).

Seven studies used severity ratings to select outcome events for statistical analysis. Two used a preventability assessment.

Table 1 – Summary of included studies

Source	Outcome event	Risk / prognostic factor identified?	Prediction tool developed?
Consensus studies			
Roten *et al.* (2010)[50]	MRP	✓	✓
Cottrell *et al.* (2013)[48]	MRP	✓	✓
Falconer *et al.* (2014)[49]	ADE & ME	✓	✓
Kaufmann *et al.* (2015)[38]	MRP	✓	
Saedder *et al.* (2016)[46]	ME	✓	✓
Hickson *et al.* (2016)[51]	MRP	✓	✓
Statistical studies			
Onder *et al.* (2010)[41]	ADR	✓	✓
Tangiisuran *et al.* (2014)[42]	ADR	✓	✓
Kiguba *et al.*(2017)[43]	ADR	✓	✓
McElnay *et al.* (1997)[44]	ADE	✓	✓
Trivalle *et al.* (2011)[45]	ADE	✓	✓
Nguyen *et al.* (2017)[47]	ME	✓	✓
Urbina *et al.* (2014)[52]	MRP	✓	✓
Bates *et al.* (1999)[76]	ADE	✓	
Van den Bemt *et al.* (2000)[77]	ADE	✓	
Blix *et al.* (2004)[78]	MRP	✓	
Evans *et al.* (2005)[79]	ADE	✓	
Johnston *et al.* (2006)[80]	ADE	✓	
Zopf *et al.* (2008)[81]	ADR	✓	
Davies *et al.* (2009)[82]	ADR	✓	
Zaal *et al.* (2010)[83]	ME*	✓	
Munoz-Torrero *et al.* (2010)[84]	ADR	✓	
Dequito *et al.* (2011)[85]	ADE	✓	
Ben-Yehuda *et al.* (2011)[86]	ME*	✓	
Beckett *et al.* (2012)[87]	ADE	✓	
O'Connor *et al.* (2012)[88]	ADR	✓	
Sikdar *et al.* (2012)[89]	ADR	✓	
Wilmer *et al.* (2015)[90]	MRP	✓	
Ashcroft *et al.* (2015)[91]	ME†	✓	
Ayalew *et al.* (2015)[72]	MRP	✓	
Lenssen *et al.* (2016)[73]	MRP	✓	

* Prescribing and transcribing errors only

† Prescribing errors only

ADR = adverse drug reaction, ADE = adverse drug event, ME = medication error, MRP = medication related problem

3.3.2.2 Prediction tool studies

Type of studies: Of the 12 prediction tool studies, five used consensus methods and seven used statistical methods (see Table 1). Of the statistical studies, six were prospective and one retrospective. All used logistic regression.

Types of participants: All 12 studies included patients on hospital medical wards / specialities, but nine also included patients from other specialities (surgery, orthopaedics, maternity, cardiology, and gynaecology). All studies included only adult patients, with four restricted to older adults (aged 65 years and over). The number of patients involved in each of the statistical studies ranged from 526 to 8,713, with a total of 18,964 patients included in the seven studies.

The majority of the studies (ten) involved research carried out in Europe (two in England and France, and one in Ireland, Scotland, Spain, Switzerland, Italy and Denmark). One study was conducted in Africa (Uganda), and one in Australasia (New Zealand).

The studies were published between 1997 and 2017.

Type of outcome measures: Four studies used MRPs as the outcome measure, three used ADRs, two used ADEs, two used MEs, and the remaining study used ADEs and MEs (see Table 1).

One study used severity ratings to select outcome events for statistical analysis. None used a preventability assessment.

3.3.3 Risk of bias in individual studies

The risk of bias was assessed for all 31 studies, and the results are separated into three sections due to the different methods used. The statistical studies are split by whether they identified prognostic factors only (18 studies), or developed prognostic models (seven studies). The six consensus studies are considered separately.

3.3.3.1 Prognostic factor studies

The results of the QUIPS review are shown in Appendix A3.6, and a summary for each of the six domains is given below.

Study participation: Eleven of the 18 studies lacked a clear description of how the study wards / sites were selected, which prevents an assessment of whether the study samples were representative of the source population. In addition, one study did not specify the dates / duration of data collection (limiting an assessment of the applicability of the results), and one did not specify whether or not patients were recruited consecutively. Although the adequacy of the sample size is dependent on the number of events per variable used for analysis[92] (as discussed in section 9.2.3), seven of the studies had fewer than 500 patients, which may affect the reliability of the findings.

Study attrition: As the included studies were linked to a defined hospital admission of limited duration this was not a significant source of bias in any of the studies.

Prognostic factor measurement: Definitions for one or more of the prognostic factors were not given in five of the 18 studies, and ten did not give details of the timing of prognostic factor data collection in relation to the patients' admission, which is a particular issue with prognostic factors that are likely to alter during the course of a patient's admission, such as the number of medicines prescribed and laboratory results. Continuous variables were categorised or cut-points used in 12 of the 18 studies, which is not recommended as it has the potential to skew the results[58]. In addition, 14 of the 18 studies did not give details of the proportion of patients with missing data, or the method used to impute any missing data.

Outcome measure: The method of measuring the outcome event was well described in 17 of the 18 studies. Regarding the validity and reliability of outcome event measurement, two studies relied on voluntary reporting systems, and one used data that were not collected for study purposes. One study used patient interviews as part of their assessment of the outcome event, which may introduce a degree of variability

(due to patients' ability to correctly recognise the outcome and attribute it to their medication, severity of illness / ability to communicate, and whether patients agree to be interviewed). In addition, one study used hospital coding data to identify the outcome event, resulting in identification of only the most severe outcomes, and one relied on twice weekly reviews by the investigators. This has the potential to result in missing data due to patient discharges, events that occur between scheduled review visits, or poor documentation in nursing / medical records.

Study confounding: All studies were exploratory and study confounding was therefore not relevant.

Statistical analysis and reporting: Five of the 18 studies gave insufficient information to judge the adequacy of the analytic strategy, and eight reported that they selected prognostic factors for inclusion in the final multivariable analysis based on initial univariable analysis. This is not recommended as it can lead to selection bias[62], and result in prognostic factors being wrongly excluded from the model[63]. In addition, four studies categorised data or used cut-points for continuous variables during the statistical analysis.

In summary, this review found that all of the prognostic factor studies have some degree of bias, which has the potential to over or underestimate the relationship between the prognostic factors and outcome events.

3.3.3.2 Prognostic model studies

The results of the CHARMS review are shown in Appendix A3.7, and a summary for each of the 11 domains is given below.

Source of data

All studies[41-45 47 52] used data collected specifically for research purposes, with six of the seven studies using a prospective cohort design, the preferred method for model development as it enables optimal measurement of prognostic factors and outcome events[60]. Onder *et al*[41] analysed data from a historical database.

Participants

Appropriate recruitment methods were used by all studies, so reducing the risk of selective sampling bias. One study used systematic random sampling[43], and all others used consecutive recruitment. All studies also clearly described the study setting and age of included participants, permitting an assessment of applicability and generalisability, although two[44 45] did not specify study dates.

It was not possible to assess the risk of selective inclusion bias for two studies as the exclusion criteria were not stated[43 47]. In addition, McElnay *et al*[44] gave the total number excluded, but no breakdown of reasons, and Tangiisuran *et al*[42] listed five exclusion criteria (self-poisoning suspected, patient transferred to another ward during weekend, admitted and discharged during weekend, died within 24 hours of admission, and medical notes not available for further investigation) but only gave details of the number who were discharged or died before end of study, which does not correspond to the initial criteria. Similarly, written consent was required for four studies[41-44], but no details given for the number, or description, of the patients who refused consent. One study did not state whether or not consent was required[45].

Outcome event

The outcome event and method of measurement were well defined in all studies. Urbina *et al*[52] used electronic alerts to identify the outcome event, ensuring reliability of data collection, but it is not clear whether all alerts were used as outcomes, or just those subsequently verified by pharmacists. All other studies used manual data collection. In one study all data were collected by one principal investigator[42], but no details of inter-rater reliability were given for the remaining studies. A consensus method to confirm and classify the outcome events was used in four studies[42 43 45 47].

A single outcome event (ADR or ADE) was used for five studies[41-43 45 47]. McElnay *et al*[44] and Urbina *et al*[52] used a combined outcome (ADR plus compliance, and MRPs respectively), but McElnay *et al*[44] did not state the frequency of the individual components, which could make comparison with other studies difficult given that prevalence may differ between populations.

Blinding, that is the assessment of the outcome event without knowledge of the predictors, occurred in one study[52] due to the use of electronically generated alerts. It was not possible to determine if blinding was used in the remaining studies.

Candidate predictors

Candidate predictors are the potential prognostic factors selected for a prognostic modelling study. The studies used a wide range of predictors, including patient demographics, clinical history, physical examination, disease characteristics, test results, and medication details, although none justified reasons for selection. Definitions, for example liver disease and heart failure, were poorly defined in all studies, as were the timing and method of predictor measurements. Some studies also used predictors with subjective definitions, which may lead to poor reproducibility[93]. For example, McElnay[44] used 'patient thinks drugs were responsible for hospitalisation' as a predictor in his final risk-score.

Categorisation of some continuous variables occurred in all studies apart from Nguyen *et al*[47], despite this being associated with reduced model reliability due to loss of statistical power, selection of spurious predictors, and overoptimistic predictive performance[60]. Ideally data on candidate predictors should be collected blindly, in terms of knowledge of the outcome event and other predictors[60 94], but it was not possible to determine if this occurred in any of the studies.

Sample size

A key consideration when assessing sample size for prognostic model studies is the number of outcome events relative to the number of variables used (which includes all candidate predictors chosen to be studied, and not just those included in the multivariable analysis, transformations for continuous predictors, and indicator variables for categorical predictors)[60]. A commonly used 'rule of thumb' is that there should be at least ten 'events per variable' (EPV)[92] but it was not possible to establish the EPV for any of the studies due to lack of information on the number of candidate predictors / variables. From the data available it appears that only three studies used an EPV of ten or more[41 43 52].

Missing data

It is recommended that prognostic studies report the frequency and type of missing data, and whether any data were imputed, or complete-case analysis used[60]. This was poorly reported in all studies. Only two studies gave details of missing data. McElnay *et al*[44] reported the number of participants who did not take part in a structured interview, but no details were given on the impact on data collection, or how missing data were handled. Kiguba *et al*[43] stated that they used the missing-assigned approach to categorise missing data for two categorical variables, but no details were given regarding how much data were missing.

Model development

All studies used logistic regression. The modelling assumptions for logistic regression include the need for: (1) observations to be independent of each other (i.e. no paired data); (2) little or no multicollinearity between predictors; and (3) predictors should be linearly related to the log odds of the outcome event[70]. In one study patients were eligible for recruitment more than once, meaning all observations were not independent; this did not appear to be accounted for in the analysis[52]. Only three studies stated that they performed checks for multicollinearity[42 44 45], of which McElnay *et al*[44] did not report the results. Similarly, only three studies reported checking linearity[43 45 47]. Nguyen *et al*[47] reported the non-linear relationship between age and outcome event, but Trivalle *et al*[45] and Kiguba *et al*[43] did not report results for their linearity analyses.

A potential source of bias in prognostic model studies is the method used to select predictors, both for initial inclusion in the multivariable modelling, and during modelling[60]. The use of univariable analysis to select predictors for modelling is known to cause prediction selection bias[60], but was used in six studies[41 42 44 45 47 52], with the remaining study not specifying how predictors were selected[43]. It is also known that inclusion of predictors that occur infrequently in the study population can lead to inaccurate results[63 95], but only two studies reported excluding predictors on this basis[42 44]. Bias also occurs when predictors are categorised[60]. Although proposed as a way to deal with non-linearity, Harrell[96] describes this as 'disastrous', as it leads to 'low predictive accuracy, serious lack of model fit, residual confounding, and overestimation of effects of remaining variables'. Despite this, categorisation was used in six studies[41-45 52].

There is no consensus on the best method to select predictors during modelling, but forward selection has been shown to be less reliable than a full model or backwards

elimination approach[60]. All methods have limitations, with substantive prior knowledge being needed for the full model approach, and a risk of omitting potentially important predictors with backward elimination[60]. It is therefore recommended that researchers clearly state their modelling approach[60], but a clear description of both the modelling strategy and criteria for predictor inclusion were only provided for two studies[42 44].

It is also recommended that shrinkage techniques are used to adjust for model overfitting[60], but only Nguyen *et al*[47] reported applying shrinkage to their model.

Model performance

It is recommended that all prognostic model studies report calibration and discrimination[60], but both measures were reported in only three studies[42 47 52]. Calibration, which shows the agreement between observed and expected predictions, was assessed in three studies[42 47 52]. Two used the Hosmer-Lemeshow test, although this has been criticised for lack of statistical power, oversensitivity in large samples, and inability to show the direction or magnitude of the miscalibration[60]. Calibration plots are preferred, but were used only by Nguyen *et al*[47]. Discrimination, which is the ability to differentiate between those who do or do not experience the outcome event, was assessed using the area under the receiver operating characteristic (ROC) curve in five studies[41 42 45 47 52]. Results ranged from 0.70 to 0.78, suggesting modest discrimination. Two studies created risk groups, therefore were able to report sensitivity and specificity[41 42], although both used a probability threshold based on study data rather than being predefined, which is associated with producing overoptimistic and biased results[60]. In addition neither study gave confidence intervals (CIs) for the sensitivity or specificity despite the recognised value of doing so[97 98]. Tangiisuran's model had the highest senstivity[42] (80%), but this was based on a detection score of one of a total score of five, hence one would expect high sensitivity at the expense of lower specificity, which was only 55%. Onder's model had poorer sensitivity (68%) but better specificity (65%)[41], which may be due to using a relatively higher detection score (four of a possible total of ten).

Model evaluation

Measuring the predictive performance of prognostic models using the same data used to develop the model, known as 'apparent performance', often overestimates the results found in a new set of patients, known as overfitting or optimism[60]. The risk is increased when using small samples, low EPV, data-driven modelling techniques, or when shrinkage techniques are not used. It is therefore recommended that performance is assessed using new data, known as model validation[60]. Internal

validation involves data-resampling, often bootstrapping or splitting the original sample. External validation involves using new data, and is therefore the preferred method[60].

Six of the studies performed model evaluation, two used external validation in a separate dataset[41 42], two used bootstrap resampling of the developmental data[45 47], and one used a non-random split of the original sample (based on the date of admission)[52]. McElnay *et al*[44] used patients from the same study site to validate their model, but did not specify whether they used a split sample, or temporal validation. Kiguba *et al*[43] did not perform model evaluation, therefore it was not possible to assess how well their model may perform in a new sample of patients.

When studies include validation in a new group of patients it is recommended that authors report the differences in the frequency / distribution of predictors and outcome events between the development and validation datasets, as it is known that case-mix may influence model performance[60]. This was reported by Tangiisuran *et al*[42], although not for all predictors. Onder *et al*[41] reported that there were differences between the two groups, but the frequencies / distributions were not given for the predictors. McElnay *et al*[44] reported that the groups were similar, but gave no further details. Urbina *et al*[52] reported the difference in outcome event prevalence only.

Onder *et al*[41], Tangiisuran *et al*[42] and Urbina *et al*[52] reported similar predictive performance in their development and validation samples. McElnay *et al*[44] reported the sensitivity and specificity of their model in the validation dataset only, but despite selecting a probability threshold to produce optimal results, concluded that their model had insufficient sensitivity (40.5%) to be a satisfactory predictor of ADEs.

Trivalle *et al*[45] and Nguyen *et al*[47] used bootstrap validation. Nguyen *et al*[47] used the results to calculate a shrinkage factor. Trivalle *et al*[45] used bootstrapping, but did not report using this to shrink the model. Instead it was used to estimate the risk of ADE for various categories of risk scores, as discussed below (under results domain).

Results

All studies reported the odds ratios and 95% CIs for predictors included in the final model, but the intercept and/or regression coefficients were not given by Onder *et al*[41] or Trivalle *et al*[45], therefore it would not be possible to use these models to estimate outcome event probabilities for individual patients.

Risk scores were developed by Onder *et al*[41], Tangiisuran *et al*[42], Trivalle *et al*[45], and Urbina *et al*[52]. Different methods were used to assign the risk scores to each variable.

Onder *et al*[41] and Urbina *et al*[52] assigned a score to each variable based on odds ratios. Although this method is commonly used to assign risk scores, it has been criticised by Moons et al[99], who advise that the use of regression coefficients is 'algebraically the only correct approach'. Urbina *et al*[52] may have created additional bias by dichotomising continuous predictors using cut-points that gave the highest sensitivity and specificity. An alternative approach recommended by Sullivan *et al*[100] is to use 'meaningful categories', for example quintiles, select a base category, then calculate the risk score based on how far each category is from the base category in terms of regression units. Trivalle *et al*[45] used regression coefficients to develop his risk score, but gave no further details on how these were used to create the score, therefore it is not possible to assess appropriateness. Tangiisuran *et al*[42] chose to assign an equal weight to each variable retained in the final model, irrespective of the regression coefficients / odds ratios. While this was chosen for simplicity, it takes no account of the strength of association between each predictor and the outcome event, therefore is unlikely to accurately reflect individual risk estimates obtained from the regression model.

As discussed above, Onder *et al*[41] and Tangiisuran *et al*[42] used their risk scores to create risk groups based on the selection of a probability threshold. Trivalle *et al*[45] and Urbina *et al*[52] did not create risk groups, but Trivalle *et al* used bootstrapping to estimate the risk of ADE for four risk score categories, for example a score of over 18 (out of maximum score of 34) corresponded to an estimated risk of 52% (95% CI 40-62%). Despite providing a useful way to interpret scores, the confidence intervals for each category were wide, and overlapped in three of the four categories, which reduces certainty and may limit clinical usefulness.

A range of predictors were included in the final prediction models**.** Although it is not possible to directly compare studies, due to lack of information on which predictors were included[43-45], and the use of unclear definitions (as stated above), Table 2 shows the predictors selected for each of the final models, and which other studies assessed the same predictors. This shows the variability in both the choice of candidate predictors, and in the associations found. Only three predictors (comorbidities, number of medicines prescribed, and age) appear in more than one model, despite many being assessed in two or more studies.

Table 2 – Predictors included in prognostic modelling studies

Predictor included in final model ✓ Selected for final model x Assessed but not included in final model	Type of adverse medication-related outcome						
	ADR			ADE		ME	MRP
	Onder *et al* [41]	Tangiisuran *et al* [42]	Kiguba *et al* [43]	McElnay *et al* [44]	Trivalle *et al* [45]	Nguyen *et al* [47]	Urbina *et al* [52]
Number of comorbid conditions / comorbidity index	✓	x	✓		x		✓
Heart failure	✓	x			x		
Liver disease	✓				x		x
Number of medicines	✓	✓	✓	x	✓	✓	✓
Previous ADR	✓	x		x			
Renal failure	✓	x			x		x
Hyperlipidaemia		✓					
Length of stay		✓		x			
Use of anti-diabetic agents / diabetes	x	✓		x			
High white cell count on admission		✓					
Age	x	x	✓	x	x	✓	✓
Gender		x	✓	x		x	x
Self-reported herbal medicine use			✓				
HIV-positive serostatus			✓				
Hospitalisation in 3 months prior to admission			✓				
Gynaecology ward			✓				
Prescribed antidepressants / depression	x			✓	x		
Prescribed digoxin				✓			
Gastrointestinal problems				✓			
Abnormal serum potassium				✓			
Thinks drugs were responsible for hospitalisation				✓			
Angina				✓			
Congestive obstructive pulmonary disease	x			✓			
Prescribed antipsychotics					✓		
Recent anticoagulation					✓		
Treatment initiated before admission						✓	
Best possible medication history available						✓	
Psycholeptics (ATC code N05)						✓	
Blood substitutes / perfusion solutions (ATC code B05)						✓	
Cardiovascular medicines (ATC group C)							✓
Hormone therapy medicines (ATC group H)							✓
Systemic anti-infective therapy (ATC group J)							✓
Sensory organ medicines (ATC group S)							✓
Medicines from ATC group V (various)*							✓
Admission to surgical versus medical ward						✓	
Hospital admission within 30 days						✓	
Admission from emergency room						✓	
Admission time (day versus night)						✓	
Admission from outside institution						✓	
MDC groups: other, nervous system, circulatory, digestive, musculoskeletal, kidney and urinary tract							✓

* Full details not given, but predominance of drugs from V03 subgroup

ADR = adverse drug reaction, ADE = adverse drug event, ME = medication error, MRP = medication related problem, HIV = human immunodeficiency virus, ATC = Anatomical Therapeutic Chemical, MDC = major diagnostic category

Interpretation and discussion

Six of the studies[41-43 45 47 52] concluded that they had produced an effective method to identify adverse medication-related outcomes, with only McElnay *et al*[44] stating that their model did not satisfactorily predict ADEs. All authors concluded that further work was needed prior to the introduction of their prediction tool into routine clinical practice. For McElnay *et al*[44] this was due to the poor predictive performance of their model[44], for other studies it was to validate the model in different populations and settings[41 43 45 47 52]. Tangiisuran *et al*[42] recommended further research prior to routine use, including the need to compare the model with clinical judgement, and to assess the usability, impact on patient safety, and associated humanistic and cost implications.

In summary, this review found that all of the prognostic modelling studies have some degree of bias, and are subject to selective reporting, which has the potential to impact both their predictive accuracy and clinical credibility.

3.3.3.3 Consensus studies

The consensus studies include one risk factor study[38], and five prediction tool studies[46 48-51]. The results of the review are summarised below, with a detailed description of each study in Appendix A3.8.

Five of the six studies used literature review to identify risk factors[38 46 49-51], but all incorporated expert opinion, either for identification[38 46 48 50 51], and/or to allocate scores[48 49]. Two studies used the Delphi method[38 46], with Kaufmann *et al*[38] using both the nominal group technique (NGT) and Delphi method, although the use of the same participants in both may have biased the results due to familiarity with the topic. The size and composition of the Delphi and NGT groups were acceptable[101], with both using representation from physicians and pharmacists. The expert groups in the remaining four studies comprised pharmacists only, with each only using staff from within the base hosptial[48-51]. The results may therefore not reflect those of a multidisciplinary group, or be generalisable to other hospitals. Three of these studies also gave limited details on the method of expert involvement / number of experts[49 50 51].

A potential source of bias for the studies that developed risk scores was the selection of the cut-off scores for risk groups. Scores were developed by Cottrell *et al*[48], Falconer *et al*[49] and Saedder *et al*[46]. Cottrell *et al* did not describe how cut-offs were selected, and both Falconer *et al* and Saedder *et al* used arbitrary cut-points. Falconer *et al* based this on staffing constraints, with the high-risk score chosen to restrict this category to the top 10% of patients. Saedder *et al* selected the score that gave the highest precision. While being a pragmatic approach, Falconer's method could be criticised for identifying patients based on workload capacity rather than actual risk, and Saedder's for the potential for overoptimistic and biased results[60].

Saedder *et al*[46] used statistical modelling to allocate risk factors scores, but may have introduced bias by the use of relatively small samples (four groups were used, ranging from 50 to 146 patients per group), and correspondingly small number of outcome events (9 to 33 per group). There were also differences in the characteristics of the groups, for example two populations were restricted to adults over the age of 65 years, whereas two included adults of all ages, and differences in the number of medicines prescribed, average number of MEs, and number of patients who experienced a ME. Finally, they used categorisation for all variables in the risk score, but it was unclear how these categories were selected.

Roten *et al*[50] and Saedder *et al*[46] reported performance measures (sensitivity and specificity), but as noted above, Saedder *et al* selected the threshold score to produce optimal performance. Roten's tool had good sensitivity (85.1%), but a correspondingly lower specificity (60.4%)[50]. Saedder's tool had better specificity (75%), but lower sensitivity (64%). Neither reported CIs. Cottrell *et al*[48] compared their algorithm with identification by traditional methods, and reported a match, but no further details were given. No details of validation testing were given by Falconer *et al*[49] or Hickson *et al*[51], although Hickson *et al* did quantify the agreement between the scores allocated by pharmacists using the tool in practice, and the per-guidance scores. No studies reported external validation.

In summary, Kaufmann's study appears to have the lowest risk of bias, but the results of all other studies may be biased due to their choice of expert group, or the method of selection for risk scores and/or groups.

3.3.4 Results of individual studies / synthesis of results

The results will be presented in two sections: prognostic factor identification, and prediction tool development studies.

3.3.4.1 Prognostic factor identification

From the 31 studies included in the review I identified 59 possible prognostic factors / groups. Table 3 provides a summary of the prognostic factor analysis, showing the proportion of studies where an association was found between the prognostic factor and outcome event. Each prognostic factor / group is shown, with the proportion of studies where an association was found, split by study method (statistical and consensus), and outcome measure. It also shows the combined proportion for all studies that included that prognostic factor or group of factors. For example, gender was associated with adverse medication-related outcomes in six of the 22 studies where it was assessed. Of these it was positively associated with the outcome event in one of the five studies of MRPs, none of the studies on MEs, and so on.

Additional information on which studies investigated each prognostic factor or group of factors is given in Table 4. This shows the prognostic factor associations by study, and is sub-divided by outcome type. For the statistical studies it gives the level of the association (none, univariable or multivariable association). For consensus studies it shows whether each prognostic factor / group was selected, or considered but not selected.

Table 3 – Summary of prognostic factor analysis

Prognostic factor	Study method / outcome event					Total
	Statistical analysis				Expert opinion	
	MRPs	MEs	ADEs	ADRs		
Demographics						
Gender	1/5*	0/2	2/7	3/8	-	**6/22**
Ethnicity	-	-	0/2	0/1	1/1	**1/4**
Non-native speaker	-	0/1	-	-	2/2	**2/3**
Age	3/5	2/3	5/7	3/9	2/4	**15/28**
Marital status	-	-	0/1	-	-	**0/1**
Patient characteristics						
Weight / height related factors	2/2	-	1/3	0/2	0/1	**3/8**
Swallowing problems	-	-	0/1	-	1/1	**1/2**
Allergy / previous ADR	-	-	1/1	1/2	2/2	**4/5**
Non-compliance with medication	-	-	-	-	2/2	**2/2**
Dependent living situation	-	-	0/2	1/3	0/1	**1/6**
Disability	-	0/1	-	0/1	-	**0/2**
Ability to sign consent form	-	0/1	-	-	-	**0/1**
Smoking status / nicotine use	-	-	0/1	0/3	-	**0/4**
Social deprivation	-	-	-	0/1	-	**0/1**
Alcohol related	-	-	0/1	1/2	-	**1/3**
Falls risk	-	-	-	0/1	0/1	**0/2**
Impaired manual skills	-	-	-	-	1/1	**1/1**
Visual impairment	-	-	0/1	-	1/1	**1/2**
Medicines related factors						
Number of medicines prescribed	5/5	3/3	4/7	7/8	3/3	**22/26**
Number of potentially inappropriate medicines prescribed	-	-	-	1/1	-	**1/1**
Cessation of medicines used before admission	-	-	1/1	-	-	**1/1**
Prescription of new medicines during / before admission	-	1/1	1/1	-	-	**2/2**
Admission details / past admissions & outpatient appointments						
Elective versus unplanned admission	0/1	-	-	-	-	**0/1**
Readmission to hospital	1/1	0/2	1/1	1/1	1/2	**4/7**
Number of past hospital admissions	-	-	0/1	-	-	**0/1**
Number of past outpatient appointments	-	-	0/1	-	1/1	**1/2**
Administrative factors						
Length of stay	0/1	2/2	2/4	5/5	-	**9/12**
Time of day prescribed	-	0/1	1/1	-	-	**1/2**
Month of stay	-	-	1/1	-	-	**1/1**
Stage of patient stay (admission / during stay / discharge)	-	1/1	-	-	-	**1/1**
Type of hospital department / speciality	3/4	1/2	3/4	3/3	-	**10/13**

Continued from previous page…

Prognostic factor	Study method / outcome event					Total
	Statistical analysis				Expert opinion	
	MRPs	MEs	ADEs	ADRs		
Medical condition						
Diagnosis / reason for admission	1/2	1/1	2/4	1/3	-	**5/10**
Comorbidities	1/2	1/1	3/4	5/6	4/4	**14/17**
Comorbidity index	2/3	1/1	3/5	4/7	1/2	**11/18**
DRG weight	1/1	-	0/1	-	-	**1/2**
Results						
Anaemia / haemoglobin	-	-	-	1/2	1/1	**2/3**
Temperature	-	-	0/1	1/1	-	**1/2**
Heart rate / blood pressure	-	-	0/1	0/1	-	**0/2**
Renal function	1/2	1/1	1/4	5/6	5/5	**13/18**
Liver function	1/1	-	0/2	2/3	2/2	**5/8**
Serum albumin	-	-	1/2	1/3	-	**2/5**
Serum amylase	-	-	-	1/1	-	**1/1**
Thyroid function	-	-	-	1/1	-	**1/1**
Hyperlipidaemia	-	-	-	1/2	-	**1/2**
White blood cell count	-	-	-	1/2	1/1	**2/3**
Platelet count	-	-	1/1	1/1	-	**2/2**
Serum potassium	-	-	1/1	1/1	1/1	**3/3**
Serum sodium	-	-	-	1/1	1/1	**2/2**
Serum calcium	-	-	-	1/1	-	**1/1**
Prothrombin time / INR	-	-	-	1/1	1/1	**2/2**
Blood glucose / HbA1c	-	-	1/1	1/1	1/1	**3/3**
Serum C-reactive protein	-	-	-	1/1	-	**1/1**
Medicine use						
Individual medicines / groups	1/2	3/3	7/7	3/4	5/5	**19/21**
'ISMP high-alert medication' / risk of harm	-	-	1/1	-	1/1	**2/2**
'Narrow therapeutic index' medicines	0/1	-	-	0/1	3/3	**3/5**
Drug interactions	-	-	-	1/1	1/1	**2/2**
Route of admin*i*stration of medication	-	2/2	2/2	-		**4/4**
Dosing frequency of medication	-	1/1	1/2	-	-	**2/3**
Drug dose (high versus low)	-	-	1/1	-	-	**1/1**

* Number of studies where an association found / total number of studies which assessed the prognostic factor or group

MRP = medication related problem, ME = medication error, ADE = adverse drug event, ADR = adverse drug reaction, DRG = Diagnosis-related group, HbA1c = Haemoglobin A1c / glycated haemoglobin test, ISMP = Institute for Safe Medication Practices

Table 4 – Detailed summary of prognostic factor analysis (showing associations by study)

Prognostic factor	Outcome event	Positive univariable correlation	Positive multivariable correlation	Selected by expert opinion / theoretical modelling	No correlation / considered by experts & not selected
Gender	MRPs	Urbina[52]			Blix[78] Wilmer[90] Lenssen[73] Ayalew[72]
	MEs				Zaal[83] Nguyen[47]
	ADEs	Beckett[87]	Evans[79]		Van den Bemt[77] Bates[76] Dequito[85] McElnay[44] Johnston[80]
	ADRs	Davies[82]	Dequito[85] Zopf[81]		O'Connor[88] Tangiisuran[42] Munoz-Torrero[84] Sikdar[89] Kiguba[43]
Ethnicity	ADEs				Bates[76] Johnston[80]
	ADRs				Tangiisuran[42]
	Expert opinion (ME/MRPs)			Falconer[49]	
Non-native speaker	MEs				Ben-Yehuda[86]
	Expert opinion (ME/MRPs)			Falconer[49] Kaufmann[38]	
Age	MRPs	Blix[78] Lenssen[73]	Urbina[52]		Wilmer[90] Ayalew[72]
	MEs	Zaal[83]	Nguyen[47]		Ashcroft[91]
	ADEs	Beckett[87] Bates[76] Johnston[80]	Van den Bemt[77] Dequito[85]		Evans[79] McElnay[44]
	ADRs	Dequito[85] Davies[82]	O'Connor[88]		Onder[41] Tangiisuran[42] Munoz-Torrero[84] Sikdar[89] Zopf[81] Kiguba[43]
	Expert opinion (ME/MRPs)			Cottrell[48] Falconer[49]	Kaufmann[38] Saedder[46]
Marital status	ADEs				McElnay[44]
Weight / height related factors	MRPs	Urbina[52] Wilmer[90]			
	ADEs		Evans[79]		Beckett[87] McElnay[44]
	ADRs				Onder[41] Munoz-Torrero[84]
	Expert opinion (ME/MRPs)				Kaufmann[38]
Swallowing problems	ADEs				McElnay[44]
	Expert opinion (ME/MRPs)			Kaufmann[38]	
Allergy / previous ADR	ADEs	McElnay[44]			
	ADRs		Onder[41]		Tangiisuran[42]
	Expert opinion (ME/MRPs)			Cottrell[48] Kaufmann[38]	

Continued from previous page…

Prognostic factor	Outcome event	Positive univariable correlation	Positive multivariable correlation	Selected by expert opinion / theoretical modelling	No correlation / considered by experts & not selected
Non-compliance with medication	Expert opinion (ME/MRPs)			Falconer[49] Kaufmann[38]	
Dependent living situation	ADEs				Bates[76] McElnay[44]
	ADRs	O'Connor[88]			Onder[41] Tangiisuran[42]
	Expert opinion (ME/MRPs)				Kaufmann[38]
Disability	MEs				Ben-Yehuda[86]
	ADRs				Tangiisuran[42]
Ability to sign consent form	MEs				Ben-Yehuda[86]
Smoking status / nicotine use	ADEs				McElnay[44]
	ADRs				Onder[41] Tangiisuran[42] Zopf[81]
Social deprivation	ADRs				Sikdar[89]
Alcohol related	ADEs				McElnay[44]
	ADRs		Zopf[81]		Tangiisuran[42]
Falls risk	ADRs				Onder[41]
	Expert opinion (ME/MRPs)				Kaufmann[38]
Impaired manual skills	Expert opinion (ME/MRPs)			Kaufmann[38]	
Visual impairment	ADEs				McElnay[44]
	Expert opinion (ME/MRPs)			Kaufmann[38]	
Number of medicines prescribed	MRPs	Wilmer[90]	Blix[78] Urbina[52] Lenssen[73] Ayalew[72]		
	MEs		Zaal[83] Ben-Yehuda[86] Nguyen[47]		
	ADEs	Bates[76]	Van den Bemt[77] Trivalle[45] Dequito[85]		Beckett[87] Evans[79] McElnay[44]
	ADRs		Dequito[85] Onder[41] O'Connor[88] Davies[82] Zopf[81] Tangiisuran[42] Kiguba[43]		Munoz-Torrero[84]
	Expert opinion (ME/MRPs)			Falconer[49] Kaufmann[38] Saedder[46]	
Number of potentially inappropriate medicines prescribed	ADRs		O'Connor[88]		
Cessation of medicines used before admission	ADEs		Van den Bemt[77]		

Continued from previous page…

Prognostic factor	Outcome event	Positive univariable correlation	Positive multivariable correlation	Selected by expert opinion / theoretical modelling	No correlation / considered by experts & not selected
Prescription of new medicines during / before admission	MEs	Nguyen[47]			
	ADEs		Van den Bemt[77]		
Elective versus unplanned admission	MRPs				Urbina[52]
Readmission to hospital	MRPs	Urbina[52]			
	MEs				Zaal[83] Nguyen[47]
	ADEs	Beckett[87]			
	ADRs		Kiguba[43]		
	Expert opinion (ME/MRPs)			Falconer[49]	Kaufmann[38]
Number of past hospital admissions	ADEs				McElnay[44]
Number of past outpatient appointments	ADEs				McElnay[44]
	Expert opinion (ME/MRPs)			Falconer[49]	
Length of stay	MRPs				Ayalew[72]
	MEs		Zaal[83] Ben-Yehuda[86]		
	ADEs	Bates[76]	Dequito[85]		Beckett[87] McElnay[44]
	ADRs	O'Connor[88] Davies[82]	Dequito[85] Tangiisuran[42] Munoz-Torrero[84]		
Time of day prescribed	MEs				Nguyen[47]
	ADEs		Beckett[87]		
Month of stay	ADEs		Beckett[87]		
Stage of patient stay (admission / during stay / discharge)	MEs		Ashcroft[91]		
Type of hospital department / speciality	MRPs	Urbina[52]	Blix[78] Wilmer[90]		Lenssen[73]
	MEs		Zaal[83]		Nguyen[47]
	ADEs		Evans[79] Bates[76] Dequito[85]		Beckett[87]
	ADRs	Davies[82]	Dequito[85] Kiguba[43]		
Diagnosis / reason for admission	MRPs		Urbina[52]		Wilmer[90]
	MEs	Ben-Yehuda[86]			
	ADEs		McElnay[44] Johnston[80]		Beckett[87] Bates[76]
	ADRs		Onder[41]		O'Connor[88] Munoz-Torrero[84]

Continued from previous page…

Prognostic factor	Outcome event	Positive univariable correlation	Positive multivariable correlation	Selected by experts / theoretical modelling	No correlation / considered by experts & not selected
Comorbidities	MRPs	Urbina[52]			Wilmer[90]
	MEs	Ben-Yehuda[86]			
	ADEs	Bates[76] McElnay[44] Trivalle[45]			Evans[79]
	ADRs	O'Connor[88] Tangiisuran[42] Munoz-Torrero[84]	Onder[41] Sikdar[89]		Kiguba[43]
	Expert opinion (ME/MRPs)			Cottrell[48] Falconer[49] Kaufmann[38] Hickson[51]	
Comorbidity index	MRPs		Urbina[52] Wilmer[90]		Ayalew[72]
	MEs		Ben-Yehuda[86]		
	ADEs	Trivalle[45]	Evans[79] Dequito[85]		Bates[76] McElnay[44]
	ADRs	Munoz-Torrero[84]	Dequito[85] Onder[41] Sikdar[89]		O'Connor[88] Tangiisuran[42] Kiguba[43]
	Expert opinion (ME/MRPs)			Kaufmann[38]	Saedder[46]
DRG weight	MRPs	Urbina[52]			
	ADEs				Bates[76]
Anaemia / haemoglobin	ADRs		Zopf[81]		Onder[41]
	Expert opinion (ME/MRPs)			Falconer[49]	
Temperature	ADEs				McElnay[44]
	ADRs		Zopf[81]		
Heart rate / blood pressure	ADEs				McElnay[44]
	ADRs				Zopf[81]
Renal function	MRPs	Urbina[52]			Wilmer[90]
	MEs	Zaal[83]			
	ADEs	Beckett[87]			Evans[79] McElnay[44] Bates[76]
	ADRs	Munoz-Torrero[84] Zopf[81]	Onder[41] O'Connor[88] Sikdar[89]		Tangiisuran[42]
	Expert opinion (ME/MRPs)			Falconer[49] Kaufmann[38] Roten[50] Hickson[51] Saedder[46]	
Liver function	MRPs	Urbina[52]			
	ADEs				Beckett[87] Bates[76]
	ADRs	Zopf[81]	Onder[41]		Munoz-Torrero[84]
	Expert opinion (ME/MRPs)			Kaufmann[38] Hickson[51]	
Serum albumin	ADEs	Bates[76]			McElnay[44]
	ADRs	Zopf[81]			Onder[41] Munoz-Torrero[84]
Serum amylase	ADRs	Zopf[81]			
Thyroid function	ADRs	Zopf[81]			
Hyperlipidaemia	ADRs		Tangiisuran[42]		Zopf[81]

Continued from previous page…

Prognostic factor	Outcome event	Positive univariable correlation	Positive multivariable correlation	Selected by experts / theoretical modelling	No correlation / considered by experts not selected
White blood cell count	ADRs		Tangiisuran[42]		Zopf[81]
	Expert opinion (ME/MRPs)			Falconer[49]	
Platelet count	ADEs		Bates[76]		
	ADRs		Zopf[81]		
Serum potassium	ADEs		McElnay[44]		
	ADRs	Zopf[81]			
	Expert opinion (ME/MRPs)			Falconer[49]	
Serum sodium	ADRs	Zopf[81]			
	Expert opinion (ME/MRPs)			Falconer[49]	
Serum calcium	ADRs	Zopf[81]			
Prothrombin time / INR	ADRs	Zopf[81]			
	Expert opinion (ME/MRPs)			Falconer[49]	
Blood glucose / HbA1c	ADEs	McElnay[44]			
	ADRs	Zopf[81]			
	Expert opinion (ME/MRPs)			Falconer[49]	
Serum C-reactive protein	ADRs	Zopf[81]			
Individual medicines / groups	MRPs		Urbina[52]		Wilmer[90]
	MEs	Nguyen[47]	Zaal[83] Ashcroft[91]		
	ADEs	Beckett[87] Dequito[85]	Van den Bemt[77] Bates[76] McElnay[44] Johnston[80] Trivalle[45]		
	ADRs		Dequito[85] Tangiisuran[42] Kiguba[43]		Onder[41]
	Expert opinion (ME/MRPs)			Cottrell[48] Falconer[49] Kaufmann[38] Roten[50] Hickson[51]	
'ISMP high-alert medication' / risk of harm	ADEs		Beckett[87]		
	Expert opinion (ME/MRPs)			Saedder[46]	
'Narrow therapeutic index' medicines	MRPs				Wilmer[90]
	ADRs				Tangiisuran[42]
	Expert opinion (ME/MRPs)			Falconer[49] Kaufmann[38] Hickson[51]	
Drug interactions	ADRs		Munoz-Torrero[84]		
	Expert opinion (ME/MRPs)			Saedder[46]	
Route of administration of medication	MEs		Zaal[83] Ashcroft[91]		
	ADEs	Beckett[87]	Evans[79]		
Dosing frequency of medication	MEs		Zaal[83]		
	ADEs	Beckett[87]			McElnay[44]
Drug dose (high versus low)	ADEs		Evans[79]		

MRP = medication related problem, ME = medication error, ADE = adverse drug event, ADR = adverse drug reaction, DRG = Diagnosis-related group, INR = International Normalised Ratio, HbA1c = Haemoglobin A1c / glycated haemoglobin test, ISMP = Institute for Safe Medication Practices

3.3.4.2 Prediction tool development studies

Prediction tools were developed in 12 of the 31 studies, five using a consensus method [46 48-51] and seven using statistical analysis[41-45 47 52]. Details for each study are given in Table 5. This shows the method of development (consensus or statistical), type of tool developed (electronic scoring algorithm, or risk model / score); prognostic / risk factors included in the prediction tool, scoring system, and results of testing / validation.

None of the 12 studies report having carried out research related to implementation or impact.

Table 5 – Summary of the prediction tool development studies

Source	Model / score	Testing / validation
Roten *et al*. (2010)[50]	Consensus derived electronic scoring algorithm based on integrated electronic systems. Six factors used to identify patients at risk of MRPs: high-risk medicines (e.g. enzyme inducers, anticoagulants); renal impairment; digoxin & low serum potassium; >3 days intravenous antibiotics; >3 days intravenous paracetamol; elderly patients with polypharmacy.	Prospective validation study of 501 patients, risk scoring compared to manual check by pharmacist (sensitivity 85.1%, specificity 60.4%).
Cottrell *et al*. (2013)[48]	Consensus derived electronic scoring algorithm based on electronic prescribing system (i.e. age & medicines related factors only). Factors included in score: age; medicine count; medicines by group (e.g. opiates, epilepsy treatment); medicines not verified; restricted supply / high-risk / non-formulary medicines; duration review needed; allergy status. Scoring varies with factor.	Patients scored by algorithm as 'high-risk' compared to those identified by 'traditional ward round' – match reported.
Falconer *et al*. (2014)[49]	Consensus derived electronic scoring algorithm based on integrated electronic systems. Thirty eight factors included in score: patient profile (e.g. age, ethnicity); patient encounter (e.g. number of previous admissions, recent readmission); clinical profile (i.e. patients with chronic disease); high-risk medication (e.g. >8 regular medicines, anticoagulants); laboratory values (e.g. renal function, serum potassium). Scoring varies with factor.	None reported.
Saedder *et al*. (2016)[46]	Consensus derived risk score based on 4 factors: renal function; number of medicines; number of high-risk medicines (subdivided into 3 risk categories); number of interactions (subdivided into 3 risk categories). Scoring varies with factor.	Sensitivity 64%, specificity 75%, area under ROC curve 0.76, (95% CI 0.62-0.89).
Hickson *et al.* (2016)[51]	Consensus derived acuity pharmaceutical assessment screening tool to assign a patient acuity (level 1, 2 or 3). Acuity assessment based on patient factors: number of decompensated organs, need for input from intensive care team, specific comorbidities (e.g. cystic fibrosis, organ transplant), use of high-risk medicine; plus a range of triggers needing pharmacist input (e.g. use of high-cost medicines, home intravenous therapy); or the need for senior pharmacist input. Acuity level allocation based on patient meeting criteria specified for each level.	Quasi-experimental service evaluation to quantify agreement among pharmacist-documented and per-guidance patient acuity level (PAL), no other testing or evaluation reported.
Onder *et al*. (2010)[41]	Statistical risk score based on 7 factors: ≥4 comorbid conditions; heart disease; liver disease; number of medicines; previous ADR; renal failure. Scoring varies with factor.	Internal validation: sensitivity 68%, specificity 65%); area under ROC curve 0.71 (95% CI 0.68-0.73) Prospective validation study of 483 patients, area under ROC curve 0.70 (95% CI 0.63-0.78). Calibration not reported.

Continued from previous page…

Source	Model / score	Testing / validation
Tangiisuran *et al*. (2014)[42]	Statistical risk score based on 5 factors: hyperlipidaemia; anti-diabetic agent; raised white blood cell count; total number of medicines; length of stay. Scoring system based on equal weighting of 1 for each factor.	Internal validation: sensitivity 80%, specificity 55%; area under ROC curve 0.74 (95% CI 0.68-0.79). Prospective validation study of 483 patients, sensitivity 84%, specificity 43%; area under ROC curve 0.73 (95% CI 0.66-0.80). Calibration reported for developmental data only.
Kiguba *et al.* (2017)[43]	Two regression models developed (for probable ADRs, and possible ADRs). Final models included: age; gender; number of conventional medicines; Charlson's comorbidity index; preadmission herbal medicines use; HIV-positive serostatus; hospitalisation in previous 3 months; gynaecology ward. Risk score calculated from regression equation.	Not performed.
McElnay *et al*. (1997)[44]	Statistical risk score based on 7 factors: antidepressants; digoxin, abnormal serum potassium; 'thinks medicines were responsible for hospitalisation'; angina; congestive obstructive pulmonary disease. Scoring system not stated.	Prospective validation study of 204 patients: sensitivity 40.5%, specificity 69%, overall accuracy 63%. Area under ROC curve and calibration not reported.
Trivalle *et al.* (2011)[45]	Statistical risk score based on 3 factors: number of medicines, antipsychotics, & recent anticoagulation. Scoring varies with factor.	Score validated by resampling technique (bootstrap): area under ROC curve 0.70 (95% CI 0.65-0.74). Calibration not reported.
Nguyen *et al.* (2017)[47]	Regression model developed. Final model included: age; number of prescribed medicines; treatment initiated before admission; best possible medication history available'; psycholeptics; blood substitutes and perfusion solutions; type of hospital admission (medical versus surgical); hospital admission within previous 30 days; admission from emergency room; admission time (day versus night); admission from an outside institution. Risk score calculated from regression equation.	Internal validation: calibration reported as 'good' but slight over-estimation of high probabilities; area under ROC curve 0.718 (95% CI: 0.689-0.748). Bootstrap validation: area under ROC curve 0.707 (95% CI not given).
Urbina *et al*. (2014)[52]	Statistical risk score based on 14 factors: age; Charlson index; number of medicines; 6 MDC groups; 5 ATC groups. Scoring varies with factor.	Internal validation: area under ROC curve 0.778 (95% CI 0.768-0.789). Prospective validation study of 4058 admissions: area under ROC curve 0.776 (95% CI 0.759-0.792). Calibration not reported.

MRP = medication related problem, ADR = adverse drug reaction, ROC = receiver operator characteristic, CI = confidence interval, HIV = human immunodeficiency virus, ATC = Anatomical Therapeutic Chemical, MDC = major diagnostic category

3.4 Discussion

Key findings

In summary, this review suggests that the existing evidence is not sufficient to definitively select prognostic factors for development of the MOAT. It also suggests that the currently available prediction tools are not suitably robust for routine clinical use, in terms of validated predictive accuracy, and/or generalisability.

Prognostic factor identification

Overall, the existing evidence is not sufficient to identify which prognostic factors should be included in a prognostic model to predict the risk of MRPs for adult patients admitted to hospital medical wards, as shown by the variation in results among studies (see Table 3). This may be due to the range in the types of studies selected for the review, including both consensus and statistical methods, and the outcome measures (MRPs, ADEs, ADRs and MEs), although as shown in Table 4, there are significant differences in results among studies assessing the same outcome event using the same research method.

This could be explained by additional differences in study design, including:

- prospective versus retrospective;
- the differences in the definitions used for each outcome type across different studies, plus further subcategorisation (e.g. Saedder *et al*[46], Zaal *et al*[83], Ashcroft *et al*[91], Ben-Yehuda *et al*[86] and Nguyen *et al*[47] all studied medication errors, but Zaal *et al* and Ben-Yehuda *et al* included only prescribing and transcribing errors, and Ashcroft restricted his study to prescribing errors);
- the range of patient groups / specialties included in the different studies (e.g. studies combining patients from different inpatient settings, such as maternity, surgery and medicine), who may have different risk factors for adverse medication-related outcomes;
- differences in the criteria used to select the outcome measures for statistical analysis, for example seven of the 25 statistical studies selected the outcome events based on severity rating, and two used a preventability rating;
- methodological limitations, as highlighted by the risk of bias assessments.

Despite the variation among studies, the results of this review are comparable with the systematic review published by Suggett *et al*[64], which reported the ten most frequently reported risk factors associated with the need for pharmaceutical interventions in a hospital setting. All ten risk factors identified by Suggett *et al* were also identified by my

review: high-risk medicines; polypharmacy; age; gender; renal function; multiple comorbidities; length of hospital stay; drug allergy; compliance-issues, and liver function. My review also identified four additional, frequently reported, predictors: diagnosis / reason for admission; readmission to hospital; weight / height related factors, and serum potassium level. Another difference between the reviews was the total number of risk factors identified, with Suggett *et al* reporting 20, whereas I identified 59. Differences may be due to the study inclusion criteria, as Suggett *et al* included literature reviews and excluded consensus studies. Suggett *et al* also excluded risk factors that could not be accessed from patient medical notes. In addition there were differences in the way risk factors were grouped, for example Suggett grouped alcohol abuse and swallowing difficulties with compliance issues, whereas I reported these separately, and I grouped 'type of hospital department / specialities' while Suggett reported these separately.

Prediction tool development studies

The studies that developed a prediction tool are subject to the same considerations and variations as above. This can be seen by the selection of different prognostic factors for models predicting the same outcome event. For example Onder *et al*[41] and Tangiisuran *et al*[42] developed prognostic models to predict ADRs in older adults, and both used logistic regression analysis, but despite there being overlap in the prognostic factors investigated, for example previous ADR, number of medicines, comorbidities, and renal function, Tangiisuran's final model only included number of medicines, whereas Onder's model included all of these.

Clinical usefulness of prediction tools is dependent on clinical credibility, accuracy, generalisability, and ideally, clinically effectiveness[55]. Each will next be discussed in turn, considering the quality, applicability and limitations of tools included in this review.

Clinical credibility covers a range of issues: the prediction outcome needs to be seen as clinically important, the tool should contain all potentially relevant predictors, and it should be straightforward to use[98]. Given that assessment of clinical credibility requires subjective judgement rather than statistical assessment, I will confine this discussion to issues related to the selection of relevant predictors and ease of use.

To enhance clinical credibility predictors should ideally be chosen based on theoretical and clinical knowledge[63], but Cottrell *et al*[48] were unable to do this as they were limited to data available in an electronic prescribing system, therefore could not include medical or laboratory-related factors. In some cases a conscious decision was taken to

exclude factors, for example Onder[41] excluded high-risk medicines, and Nguyen *et al*[47] excluded laboratory results, diagnostic categories or comorbidities. While there may be legitimate reasons, it may make users question whether the model adequately assesses all sources of risk.

Regarding ease of use, automated electronic scoring systems permit the use of complex and unambiguous scoring systems, but potential issues with the manual scores include:

- unclear definitions used for predictors in all studies;
- overlap in the categories for 'number of medicines' ('≤5' and '5-7') in the score developed by Onder *et al*[41];
- no recommendation provided regarding required course of action, i.e. no risk groups created (Kiguba *et al*[43], Trivalle *et al*[45], Nguyen *et al*[47], Urbina *et al*[52]);
- the need for users to:
    - categorise the risks of harm and interaction for each medicine (Saedder[46]);
    - calculate the Charlson index (Urbina *et al*[52] and Kiguba *et al*[43]);
    - categorise medicines and diagnostic category using Anatomical Therapeutic Chemical (ATC) and major diagnostic category (MDC) classifications respectively (Urbina *et al*[52]);
    - calculate the risk score using the regression equation, i.e. no simplified scoring system developed (Kiguba *et al*[43] and Nguyen *et al*[47]).

An additional issue with Tangiisuran's tool[42] was the inclusion of 'length of stay' as a predictor, despite this not being known prior to discharge.

Predictive accuracy is a product of robust methodology at all stages of prediction tool development[60]. It was not possible to fully assess the predictive performance of all studies as four did not report performance measures[43 48 49 51], but the results of the others suggest modest overall performance. Eight studies reported performance in terms of the sensitivity, specificity and/or area under the receiver operating characteristic (ROC) curve[41 42 44-47 50 52]. Excluding McElnay *et al*[44], who concluded that their model had insufficient sensitivity, the remaining prediction tools had adequate discrimination capacity[102] (area under the ROC curve 0.70-0.78), reasonable sensitivity (64-85%), but poor-moderate ability to correctly identify those without the outcome event (specificity 43-75%). This means the risk-scores may fail to identify high-risk patients, while incorrectly categorising others as high-risk, so reducing the ability of the risk-score to manage workload efficiently. Model calibration (a measure of agreement

between predictions and observed outcomes[94]), was only reported for two studies[42 47]. While calibration is likely to be good for model development studies (as the model is optimised for the developmental data), the reporting of calibration is considered 'fundamental' for validation studies[94]. Despite this, neither of the studies that carried out external validation[41 42] reported model calibration in the validation dataset, despite one reporting it for model development[42].

Validation studies for five of the prediction tools have subsequently been published, with all finding that the tools performed less well than in the original studies[88 103-106]. Three studies have validated Onder's score[88 103 104]. O'Connor *et al*[88] concluded that Onder's score incorrectly classified 38% of patients as low-risk (area under ROC curve 0.62, 95% CI 0.57-0.68), and Petrovic *et al*[103] found that the score had 'fair' accuracy (area under ROC curve 0.64, 95% CI 0.55-0.74), good sensitivity, but very poor specificity across a range of subpopulations. Stevenson[104] compared the performance of Onder's[41] model with Tangiisuran's[42] and Trivalle's models[45], and concluded that there was poor agreement among scores, with only four patients (of 270) being categorised as high-risk by all three risk-scores. Of the consensus studies, Falconer *et al* published the results of a validation study of her original risk-score[105]. This found no significant differences in prescribing errors among the risk groups. Bonnerup *et al*[106] have also published a study validating Saedder's score[46]. Although the sensitivity and specificity are not stated in the paper, it is possible to calculate them from the results (sensitivity 57%, specificity 85%), suggesting better specificity but lower sensitivity than the development study. The tool failed to correctly identify 22 (42%) of 52 patients who experienced a ME, so one could argue that the sensitivity may be too low to instil clinical confidence in users. One option could be to lower the detection limit, as this may improve the sensitivity, but this would also reduce the specificity, so increasing the number of patients who would require a medication review. Using the original detection limit, 37% of patients in Bonnerup's study required medication review, which they suggest is in accordance with other prediction tools, so given the high outcome prevalence in their study (50.5%), it may not be practical to aim for 100% sensitivity, as this could result in the majority of patients requiring review. An alternative prediction outcome, based on severity or clinical significance, may be more pragmatic and relevant to practice. Although this is not a conclusion the authors reach, it is supported by the finding that only 29 (60%) of 48 recommendations made by pharmacists for the high-risk category patients were implemented by the hospital physician, which could suggest the remaining ME were not considered to be clinically significant.

Generalisability includes whether the tool is likely to be applicable to other settings. This depends on: the source of data used for the study (that is, whether the sample adequately represents the true population); whether the patient characteristics of the study sample reflect patients at different sites; whether the tool is up-to-date / reflects practice at the new site; and whether the outcome event and predictors are relevant to new sites, and measured in similar ways[60].

Potential issues with existing prediction tools include:

- the need for integrated electronic information systems[48-50 52];
- potentially differences in patient characteristics due to country where study conducted:
    - Kiguba's study took place in Uganda, where the HIV-positive serostatus and use of herbal medicines (two predictors in the tool) are potentially higher than in European countries;
    - high-risk medicines were included in ten tools[42 44-52], but usage patterns may differ among countries;
- age of participants (Onder *et al*[41], Tangiisuran *et al*[42], McElnay *et al*[44] and Trivalle *et al*[45]) developed tools for older adults, which may not be transferable to younger patients;
- four studies excluded patients who were unable to provide written consent[41-44], which may mean acutely unwell / confused patients may be inadequately represented;
- Urbina *et al*[52] used an integrated electronic warning system unique to their study site to identify outcome events, and therefore results may not be reproducible.

Clinical effectiveness in improving decision making and patient outcomes has not been adequately assessed in any of the studies. Nguyen *et al*[47] compared strategies for pharmaceutical intervention, comparing the model predictions with patient age to identify patients in need of intervention, and found that pharmacists were more likely to identify MEs using the model than age alone, although in practice age would not be used alone to prioritise patients.

In summary, none of the prediction tools included in this review have evidence for sufficient predictive accuracy and/or generalisability to recommend them for routine use outside of the site where they were developed. I am aware that Hickson's tool[51] is routinely used across hospitals in Manchester, England, but due to acknowledged limitations, the team who developed it are in the process of developing a more

comprehensive prioritisation-tool as part of an National Institute for Health Research (NIHR) funded study[107].

Strengths and limitations

Strengths of this review include the use of reporting guidelines for systematic reviews[66] and prognostic modelling studies[60], and the use of recognised tools to assess the risk of bias for prognostic studies[60 69]. I also formulated a clear review question, used a range of databases for the initial search, and followed this with weekly alerts to identify subsequent publications.

There are a number of limitations, the key one being the lack of independent review during study selection, data extraction, and critical appraisal[68]. Ideally I would also have included grey literature, and studies published in languages other than English, but these were excluded for pragmatic reasons. The Cochrane Library could also have been used as an additional information source. It was also not possible to assess the likelihood of publication bias as there is currently no requirement to register prognostic studies.

Implications for future research

The findings of this literature review led me to the following recommendations for development of the MOAT:

- robust methodology, guided by the PROGRESS recommendations[53 55 58 59], should be used at all stages of development to improve predictive accuracy and credibility;
- the selection of prognostic factors should be guided by expert opinion in addition to the findings of the literature search (to build on existing research, and enhance clinical credibility);
- the prognostic factors should be clearly defined, measureable, and routinely used in clinical practice (to prevent the need for additional measurements, calculations or complex categorisation prior to use of the MOAT);
- a clinically relevant outcome event, based on severity / clinical significance should be used (to ensure workload efficiency);
- outcome identification should be reliably assessed, involving inter-rater reliability assessment among data collectors, and a consensus method to validate and classify the outcome event;
- high-risk medicines should be grouped rather than modelled individually to increase generalisability;

- model performance should be reported as the calibration and discrimination, and shrinkage should be applied to account for model overfitting;
- use of the MOAT should not be predicated on electronic capability such as integrated electronic systems, as this would limit generalisability;
- to enhance ease of use; the MOAT should be presented as a simplified scoring system to avoid the need for complex calculations;
- the MOAT should provide users with recommendations regarding required course of action (i.e. risk groups);
- validation, and impact and implementation studies, should be undertaken following MOAT development.

3.5 Conclusion

The literature review has permitted the identification of 59 possible prognostic factors / groups that are associated with adverse medication-related outcomes in hospitalised adult patients. It has also permitted a review of the quality, applicability and limitations of existing prediction tools for this outcome, and highlighted the need to follow sound methodological principles in order to minimise the risk of bias, and optimise predictive accuracy.

In chapter 2 of the thesis I described a gap in the evidence base, the need for a methodologically sound prognostic model to target hospital patients most in need of pharmacists' input, based on their risk of MRPs. This literature review has clarified how this gap should be addressed, and steps that should be taken to ensure the MOAT is usable in clinical practice. This led to development of aim and objectives, which will be presented in the next chapter.

Chapter 4: Aim and objectives

The aim of this study was to use prognostic modelling to develop a prediction tool, the Medicines Optimisation Assessment Tool (MOAT^{TM}), to identify adult patients at highest risk of moderate or severe preventable medication related problems (MRPs) during admission to a medical ward in the United Kingdom, irrespective of age. It is proposed that the MOAT could be used to increase the efficiency of hospital pharmacy services, reduce risks and improve patient outcomes.

MRPs were chosen as the prediction outcome to permit the MOAT to identify patients in need of medicines optimisation, not simply those at risk of medication-related harm. Moderate or severe MRPs were chosen to ensure the MOAT targets patients most in need of pharmacists' input. This is to ensure the MOAT is clinically relevant and feasible to implement in terms of workload for pharmacists. Similarly, preventable MRPs were chosen to ensure the MOAT identifies patients with MRPs that are amenable to pharmacist intervention. No other methodologically sound prognostic model to target hospital patients based on their risk of moderate or severe preventable MRPs currently exists.

The objectives were to:

- identify potential prognostic factors based on evidence from previous research, expert opinion, and suitability in terms of methodological requirements for prognostic models;
- use prognostic modelling to develop a decision aid (the MOAT) for use in clinical practice to allocate patients to risk groups;
- assess predictive performance of the MOAT using calibration, discrimination, sensitivity, and specificity;
- review the MOAT's content validity, feasibility of use, potential efficiency savings, and the potential clinical risk associated with false negative predictions.

The first stage in developing the MOAT was to select the potential prognostic factors, also known as candidate predictors. This will be discussed in more detail in the next chapter.

Chapter 5: Selection of candidate predictors

5.1 Introduction

Potential prognostic factors, also known as candidate predictors, are the variables that predict the outcome event of a prognostic model. These can include patient demographics, clinical history, physical examination, disease characteristics, test results, and treatments used[93].

When developing a prognostic model it is necessary to limit the number of candidate predictors. Selecting too many can result in 'overfitting'/'optimism' (type I errors) leading to an overestimation of the predictive performance of the model, and the selection of spurious predictors[94]. It can also cause 'underfitting' (type II errors), which increases the risk that important predictors are not included in the model[94]. Both can lead to poor performance when the model is used in a new set of subjects[94]. One method to reduce the number of candidate predictors is to base the selection on the univariable association between each predictor and the outcome event. This is not recommended as it results in overfitting due to selection bias[62], and can lead to predictors being wrongly excluded from the model due to the fact that the association may only become significant after adjustment for the other predictors[63]. It is therefore recommended that the candidate predictors are selected *a priori*[58 95].

The choice of potential candidate predictors is ultimately guided by the intended use of the model, but various recommendations have been made:

- predictors already reported as prognostic should be included[62 63];
- the selection should be informed by clinical understanding (i.e. expert opinion) to ensure the list is comprehensive and clinically relevant[56 63];
- where predictors are highly correlated (e.g. weight and body mass index), only one should be selected[63];
- use of predictors that occur infrequently can lead to inaccurate results[63 95];
- candidate predictors should be:
    - available at the time when the model is intended to be used[93];
    - clearly defined, standardised, and reproducible (to enhance generalisability and applicability of study results to practice)[93];
    - have minimal measurement error (as this may dilute their prognostic value)[60].

An additional consideration for this study was the selection of predictors that form part of standard clinical datasets. This was to increase the reliability of the data and minimise the potential for missing data, and to enable the Medicines Optimisation Assessment Tool (MOAT^{TM}) to be readily incorporated into clinical practice without the need for additional tests / measurements.

These recommendations therefore provide three key elements to consider when selecting candidate predictors: (1) evidence from previous research; (2) the consensus of clinical experts; and (3) suitability in terms of methodological requirements.

The aim of the work presented in this chapter was to pre-select the candidate predictors for development of the MOAT. The objectives were to:

- review the potential predictors identified by the literature review to identify those reported as prognostic;
- carry out an expert survey of healthcare professionals and patient / public representatives to obtain clinical understanding and lay views on potential predictors;
- review each potential predictor, based on the above recommendations (theoretical knowledge, clinical understanding, and methodological considerations) to pre-select candidate predictors for the MOAT.

5.2 Methods

The selection of candidate predictors was split into three stages: the selection of predictors for the expert survey, expert survey, and final candidate predictor selection. Each of these is described in turn.

A summary of the methods used to operationalise the candidate predictor selection recommendations (listed in the introduction, section 5.1) is given in Table 6.

Table 6 – Assessment of predictors against selection recommendations

Methodological recommendation	Method of operationalision
Predictors reported as prognostic should be included[62 63]	Literature review performed and simple count used to identify predictors. Strength of evidence categorised as 'low' if association found in 33% or fewer studies (that investigated the predictor), 'moderate' if higher than 33% but fewer than 66%, and 'high' if higher than 66%.
Selection of predictors should be informed by clinical understanding[56 63]	Expert survey of healthcare professionals and patient / public representatives to obtain clinical understanding and lay views on potential predictors.
If predictors are highly correlated only one should be selected[63]	Where high interdependency anticipated, one predictor excluded based on level of evidence / other methodological recommendations.
Predictors that occur infrequently can lead to inaccurate results[63 95]	Predictor excluded if estimated occurrence <10% patients*.
Predictors should be available at the time model intended to be used[93]	Predictor excluded if data not available on day patient admitted to hospital.
Predictors should be clearly defined, standardised, and reproducible[93]	Predictor excluded if definition subjective and/or subjective measurement scale used.
Predictors should have minimal measurement error[60]	Predictor excluded if low reliability anticipated with test-retest, intra-rater and/or inter-rater measurements.
Predictors should form part of standard clinical datasets	Predictor excluded if not included in standard medical records and/or estimated data availability <50% patients†.

* 5% used in previous prognostic model studies[42 44 45], but 10% selected to allow for estimation error (as review based on personal clinical experience / knowledge)

† Based on Steyerberg's[70] recommendation that predictors with more than 50% missing data generally mistrusted

5.2.1 Selection of predictors for the expert survey

The selection of predictors for the expert survey involved two stages: (1) identification of predictors previously reported as prognostic; and (2) an assessment of suitablilty in terms of methodological requirements.

To establish which predictors have previously been reported as prognostic I reviewed 31 published studies (see literature review, chapter 3). I then made an initial assessment of each predictor against the remaining methodological recommendations for candidate predictors using personal clinical experience / knowledge, operationalising each recommendation as summarised in Table 6. Predictors were excluded if they did not meet all of the methodological requirements. All remaining predictors were included in the expert survey.

5.2.2 Expert survey

I developed an electronic survey, comprising both open and closed questions, to obtain expert opinion on: (1) the perceived importance / clinical relevance of the proposed predictors; and (2) other potential predictors. During development I circulated the survey to members of the MOAT project steering group (comprising PhD supervisors; clinical supervisors; patient / public representatives) to obtain comments on readability, layout, and content, and amended the survey accordingly. See Appendix A5.1 for the final version of the survey. The survey was administered during April-June 2016, with target subjects comprising both healthcare professionals and patient / public representatives. Invitations to participate were shared through the following fora / networks:

- Royal Pharmaceutical Society of Great Britain Research and Evaluation Network;
- Eastern Academic Health Science Network;
- CHAIN (Contact, Help, Advice and Information) Network;
- Medication Safety Officers Network for England;
- East of England Chief Pharmacists' Group.

It was also emailed directly to:

- Medication Safety Team, NHS England;
- other researchers in the field (identified during the literature review);
- pharmacists, consultants and senior nurses at the study sites;
- members of the MOAT project steering group.

All respondents were also requested to share the survey further within their networks / organisations.

Respondents were asked to rate each potential predictor using five Likert options (very important, important, 50:50, less important, not important). The median and interquartile range was calculated for each predictor to establish central tendency and variability, treating responses as ordinal data.

Respondents were also asked whether the MOAT should include or exclude an assessment of topical medicines. This was to establish if the MOAT should be limited to assessing risk associated with medicines taken internally, specifically by mouth, injection, inhalation, rectally or vaginally, excluding topical medicines such as creams and eye drops. Respondents were also asked to suggest other predictors that should be included.

5.2.3 Final candidate predictor selection

I made the final selection of candidate predictors following an assessment of the potential predictors included in the expert survey, and those suggested by survey respondents. I assessed each predictor against the recommendations for candidate predictors: the evidence from previous research (where available), clinical understanding and lay views (in terms of survey scores for perceived importance / clinical relevance, or number of times additional predictors suggested by respondents), and the remaining methodological recommendations for candidate predictors as summarised in Table 6 (see below for alteration to method used to assess if predictors were part of standard clinical datasets[ii]). I also considered each of the predictors suggested by survey respondents in terms of appropriateness for identifying patients at risk of MRPs during admission rather than post-discharge. This was based on personal clinical understanding.

To review the evidence from previous research I performed a simple count of the studies that identified a predictor as prognostic (a recognised method to identify key predictors[70]). As there is no evidence base / precedent to categorise strength of evidence for predictors, I used a pragmatic approach, namely to categorise the evidence as 'low', 'moderate' or 'high', which is consistent with the categorisation used in the Quality in Prognostic Studies tool[69] to assess methodological bias within prognostic studies. Evidence was categorised as low if an association was found in 33% or fewer (of the studies that investigated the predictor), 'moderate' if higher than 33% but fewer than 66%, and 'high' if higher than 66%.

In summary, predictors were selected for use in development of the MOAT if they had:

- moderate or high strength of evidence, and/or were categorised as important or very important by survey respondents;
- met all of the methodological requirements;
- were related to the risk of MRPs during admission, rather than post-discharge.

[ii] To identify if predictor data were routinely collected as part of standard clinical datasets I reviewed the standard admission proformas used at each study site. Where data were included on the proformas I then reviewed the medical records of 84 patients (50 from Hospital A and 34 from Hospital B) to establish the actual, rather than perceived frequency of data availability.

5.3 Results

The results will be presented in three sections: selection of predictors for the expert survey; expert survey, and final candidate predictor selection. Each of these is described in turn.

5.3.1 Selection of predictors for the expert survey

The literature review identified 59 possible predictors (Table 3, chapter 3). Following a review of the methodological requirements, 26 predictors were selected for inclusion in the expert survey. The predictors that were not selected are summarised in Table 7, with reason(s).

Social deprivation was initially excluded from the expert survey, but members of the MOAT project steering group suggested it may be a potentially significant predictor. I therefore decided to include it on the basis that while it is not routinely recorded, it can be calculated from postcode data.

The 27 predictors included in the expert survey are listed in section 5.3.2.

Table 7 – Potential predictors excluded from the expert survey with reason(s)

Predictor	Reason(s) for non-selection						
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Other (see details)
Non-native speaker				✓	✓	✓	
Marital status							Unlikely to be associated with MRPs during hospital stay
Weight / height related factors						✓	
Non-compliance with medication				✓	✓	✓	
Disability				✓	✓	✓	
Ability to sign consent form				✓	✓	✓	
Smoking status / nicotine use				✓	✓	✓	
Social deprivation						✓	
Alcohol related				✓	✓	✓	
Falls risk				✓	✓	✓	
Impaired manual skills				✓	✓	✓	
Visual impairment				✓	✓	✓	
Cessation of medicines used before admission				✓	✓	✓	
Prescription of new medicines during / before admission				✓	✓	✓	
Length of stay			✓				
Time of day prescribed						✓	Not patient specific
Month of stay							Not patient specific
Stage of patient stay (admission / during stay / discharge)							MOAT not intended to target stage of pharmacist input
Comorbidity index	✓					✓	Plan to use ‘number of comorbidities
DRG weight	✓					✓	Plan to use diagnostic categories

Continued from previous page…

Predictor	Reason(s) for non-selection						
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Other (see details)
Anaemia / haemoglobin	✓						
Temperature	✓						
Heart rate / blood pressure	✓						
Serum amylase						✓	
Thyroid function						✓	
Serum calcium	✓						
Prothrombin time / INR	✓						
Blood glucose / HbA1c	✓					✓	
Serum C-reactive protein	✓						
'ISMP high-alert medication' / risk of harm							Plan to use alternative method to categorise high-risk medicine use.
'Narrow therapeutic index' medicines							Plan to use alternative method to categorise high-risk medicine use.
Drug interactions				✓	✓		
Drug dose (high versus low)				✓	✓		

* Estimated occurrence <10%

† Not included in in standard medical records and/or estimated that data available for <50% patients

DRG = Diagnosis-related group, INR = International Normalised Ratio, HbA1c = Haemoglobin A1c / glycated haemoglobin test, ISMP = Institute for Safe Medication Practices

5.3.2 Expert survey

A total of 247 responses were received. Table 8 summarises the professional role of the 237 respondents who answered this question.

Table 8 – Current role of survey respondents

Current role	Number	Percentage
Pharmacist / member of the pharmacy team	178	75.2
Doctor	31	13.1
Nurse	10	4.2
Academic (no other professional role stated)	10	4.2
Patient or public representative	6	2.5
Other healthcare professional	2	0.8

Due to the 'infinite' target population it is not possible to determine a response rate.

Perceived importance / clinical relevance of predictors

The survey found that the majority of predictors (23/27) were considered 'important' or 'very important'. Details are given in Table 9.

Table 9 – Categorisation of the perceived importance of the proposed predictors as determined by median response

Predictor	Survey results	
	Median response*	Interquartile range
Renal function	1	0
Liver function	1	1
Age	1	1
Comorbidities	1	1
Allergies	1	1
Swallowing problems	1	1
Number of medicines prescribed	1	1
Number of potentially inappropriate medicines prescribed	1	1
Type of medicine prescribed	1.5	1
Serum sodium level	2	1
Serum potassium level	2	1
Platelet count	2	1
Serum albumin level	2	1
White blood cell count	2	2
Diagnosis / reason for admission	2	1
Type of hospital department / speciality	2	1
Readmission to hospital within 30 days	2	1
Number of hospital admissions within 6 months	2	1
Elective versus unplanned admission	2	1
Route of administration of medication	2	1
Dosing frequency of medication	2	1
Social deprivation	2	1
Dependent living situation	2	1
Ethnicity	3	2
Hyperlipidaemia	3	2
Number of outpatient appointments within 6 months	3	1
Gender	4	1

* Likert responses allocated ordinal numbers, 1= very important, 2=important, 3=50:50, 4=less important, 5=not important

Inclusion of topical medicines

A total of 247 responses were received to the question related to the inclusion / exclusion of topical medicines in the development of the MOAT: 82 (33%) answered 'yes' (it is acceptable to exclude topical medicines), 129 (52%) answered no, and 36 (15%) were 'unsure'. Below are all comments received related to the exclusion of topical medicines.

- "Deciding to leave out eye drops is not a good idea - should not exclude patients with glaucoma."
- "Not necessarily a risk factor but by not including topical medicines how will you identify use of patches especially opiate for pain."
- "Opioid patches are applied topically to the skin and are an example of a medicine that should be included, so there needs to be greater consideration about exclusions / inclusions."
- "Unclear whether patches would be included, but I believe they should be."
- "Patches, like fentanyl patches are associated with considerable patient safety issues, thus it is not appropriate to exclude them from the study."

Additional predictors suggested

One hundred and fifty respondents suggested additional predictors they felt should be considered for inclusion in the MOAT. These were categorised into 59 predictors / groups as listed in Table 10.

Table 10 – Additional predictors suggested by survey respondents

Predictor	Number of suggestions
Medicine related	
Over the counter / herbal medicine use	5
Length of time on medicine / newly prescribed	5
Medication Regimen Complexity Index / "complex" medication regimen	2
Irregular dose and administration / unusual dosage regimens	2
Anticholinergic burden	2
Medicine use ‘off label’	1
Homecare provided medicines	1
Length and appropriateness of antibiotic treatment	1
Use of an antidote e.g. naloxone, vitamin K	1
Constituents in formulations that may be pharmacologically active	1
Medicines or combination of medicines that predispose falls	1
Patient related	
Dementia / cognitive function / mental capacity / mental health status / confusion / delirium	34
Adherence / compliance	17
Physical / sensory impairment	14
Patient health beliefs / behaviours	11
Compliance aid	11
Frailty score	10
Language barrier	9
Self-care for medicines / whether patient / family / carer is responsible for medicines	9
Carer status	7
Intellectual disability / learning difficulty	6
Weight (obese and anorexia)	6
Poor health literacy	5
Patient education level / literacy	5
Recreational drugs / substance misuse	4
End of life care	3
Patient / carer level of knowledge / patient baseline understanding of disease state / medication	3
Alcohol use / misuse	3
Nil by mouth / enteral tube	3
Social / cultural issues	3

Continued from previous page…

Predictor	**Number of suggestions**
Overdose risk / previous overdose / misuse of medication	2
Falls risk	2
Housing status (e.g. homeless)	2
Activities of Daily Living score / functional level	2
Medicines related admission	2
Social-related admission	1
Capacity as defined by Mental Capacity Act	1
Member of travelling community	1
Elderly living alone	1
Decanting of medicines occurring	1
Significant weight changes	1
Bariatric patient	1
Nutritional status	1
Smoking status	1
Transgender	1
Housebound	1
Disability	1
Pain score	1
Venous access patient / type of cannula	1
Having received antibiotics in last 3 months	1
Frequency of GP contact	1
Identifying if patient is on a risk register with general practitioner	1
Pregnancy / breastfeeding	1
Requirement to manipulate the medicine before administration	1
Critical care admission	1
System / process related	
Staffing levels on ward / hospital	5
Communication problems across interfaces	4
The days admitted / time of year	3
Number of patient transfers across wards	1

Additional suggestions were received regarding how ‘diagnosis’ and ‘type of medicine prescribed’ should be categorised for the analysis. Suggestions for diagnoses included diabetes, Parkinson’s disease, immunodeficiency, and sepsis. Suggestions for analysis of medicines included ‘drugs with narrow therapeutic index’, ‘intravenous antibiotics’, ‘anticoagulants’, ‘insulin’, and ‘antiepileptics’.

5.3.3 Final candidate predictor selection

Predictors included in the expert survey

The review of predictors included in the expert survey is summarised in Table 11, together with those pre-selected as candidate predictors.

Table 11 – Review of candidate predictors included in expert survey

Predictor	Strength of published evidence*	Survey results		Low correlation with other predictor(s)	Estimated occurrence ≥ 10%	Available when model intended to be used	Clearly defined / reproducible	Minimal measurement error	Part of standard clinical datasets / reliably recorded+	Selected as a candidate predictor
		Median response+	Interquartile range							
Renal function	High	1	0	✓	✓	✓	✓	✓	✓	✓
Liver function	Mod	1	1	✓	✓	✓	✓	✓	✓	✓
Age	Mod	1	1	✓	✓	✓	✓	✓	✓	✓
Comorbidities	High	1	1	✓	✓	✓	✓	✓	✓	✓
Allergies	High	1	1	✓	✓	✓	✓	✓	✓	✓
Swallowing problems	Mod	1	1	✓	✓	✓	X	X	X	X
Number of medicines prescribed	High	1	1	✓	✓	✓	✓	✓	✓	✓
Number of potentially inappropriate medicines prescribed	High	1	1	✓	✓	✓	✓	✓	X	X
Type of medicine prescribed	High	1.5	1	✓	✓	✓	✓	✓	✓	✓
Serum sodium level	High	2	1	✓	✓	✓	✓	✓	✓	✓
Serum potassium level	High	2	1	✓	✓	✓	✓	✓	✓	✓
Platelet count	High	2	1	✓	✓	✓	✓	✓	✓	✓
Serum albumin level	Mod	2	1	✓	✓	✓	✓	✓	✓	✓
White blood cell count	High	2	2	✓	✓	✓	✓	✓	✓	✓
Diagnosis / reason for admission	Mod	2	1	✓	✓	✓	✓	✓	✓	✓
Type of hospital department / speciality	High	2	1	X	✓	✓	✓	✓	✓	X

Continued from previous page…

Predictor	Strength of published evidence*	Survey results		Low correlation with other predictor(s)	Estimated occurrence ≥ 10%	Available when model intended to be used	Clearly defined / reproducible	Minimal measurement error	Part of standard clinical datasets / reliably recorded‡	Selected as a candidate predictor
		Median response†	Interquartile range							
Readmission to hospital within 30 days	Mod	2	1	X	✓	✓	✓	✓	X	X
Number of hospital admissions within 6 months	Low	2	1	✓	✓	✓	✓	✓	✓	✓
Elective versus unplanned admission	Mod	2	1	✓	X	✓	✓	✓	✓	X
Route of administration of medication	High	2	1	✓	✓	✓	✓	✓	✓	✓
Dosing frequency of medication	High	2	1	X	✓	✓	✓	✓	✓	X
Social deprivation	Low	2	1	✓	✓	✓	✓	✓	X	✓§
Dependent living situation	Low	2	1	✓	✓	✓	X	X	X	X
Ethnicity	Low	3	2	✓	✓	✓	✓	✓	✓	X
Hyperlipidaemia	Mod	3	2	✓	✓	✓	✓	✓	X	X
Number of outpatient appointments within 6 months	Low	3	1	X	✓	✓	✓	✓	✓	X
Gender	Low	4	1	✓	✓	✓	✓	✓	✓	X

* Strength of evidence categorised as 'Low' association found in ≤33% published studies, 'Mod' (moderate) >33% and <66%, 'High' ≥66% (see Table 3, chapter 3 for a detailed breakdown)

† Likert responses allocated ordinal numbers, 1= very important, 2=important, 3=50:50, 4=less important, 5=not important

‡ Included in standard admission proforma at study sites and/or data available for ≥50% patients (based on a review of the patient records of 84 patients, 50 from Hospital A and 34 from Hospital B)

§ Not routinely recorded in medical records, but can be calculated from postcode

Additional predictors suggested by survey respondents

The review of additional predictors suggested by survey respondents is summarised in Table 12. Two of the 59 predictors, dementia and weight, were selected as candidate predictors. Dementia was included as dementia / cognitive function received a high number of suggestions (34), and dementia meets the remaining methodological requirements. I had previously excluded weight based on estimated data availability, but the further assessment found that it would be possible to calculate the body mass index for 62 of the 84 patients reviewed (74%) therefore this was also selected as a candidate predictor.

Table 12 – Review of candidate predictors suggested by survey respondents

Predictor	Reason(s) for non-selection							
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Related to problems encountered by patients in primary care	Other / comments
Over the counter / herbal medicine use				✓	✓	✓		
Length of time on medicine / newly prescribed				✓	✓	✓		
Medication Regimen Complexity Index						✓	✓	
Irregular dose and administration				✓	✓		✓	
Anticholinergic burden						✓		
Medicine use ‘off label’						✓		
Homecare provided medicines		✓					✓	
Length and appropriateness of antibiotic treatment								Antibiotics to be analysed as high-risk medicines
Use of an antidote e.g. naloxone, vitamin K								Sign of MRP rather than predictor
Constituents in formulations that may be pharmacologically active	✓							
Medicines or combination of medicines that predispose falls				✓		✓	✓	

Continued from previous page…

Predictor	Reason(s) for non-selection							
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Related to problems encountered by patients in primary care	Other / comments
Dementia								INCLUDE
Cognitive function / mental capacity / mental health status / confusion / delirium				✓	✓	✓		
Adherence / compliance				✓		✓	✓	
Physical / sensory impairment				✓		✓	✓	
Patient health beliefs / behaviours				✓		✓	✓	
Compliance aid						✓	✓	
Frailty score						✓	✓	
Language barrier				✓		✓	✓	
Self-care for medicines / whether patient / family / carer is responsible for medicines				✓		✓	✓	
Carer status				✓		✓	✓	
Intellectual disability / learning difficulty		✓		✓		✓	✓	
Poor health literacy				✓		✓	✓	
Weight (obese and anorexia)								INCLUDE
Patient education level / literacy				✓		✓	✓	
Recreational drugs / substance misuse		✓		✓		✓	✓	
End of life care		✓						
Patient / carer level of knowledge / patient baseline understanding of disease state / medication				✓		✓	✓	
Alcohol use / misuse				✓			✓	
Nil by mouth / enteral tube						✓		
Social / cultural issues				✓		✓	✓	

Continued from previous page…

Predictor	Reason(s) for non-selection							
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Related to problems encountered by patients in primary care	Other / comments
Overdose risk / previous overdose / misuse of medication				✓		✓	✓	
Falls risk				✓		✓	✓	
Housing status (e.g. homeless)		✓		✓		✓	✓	
Activities of Daily Living score / functional level						✓	✓	
Medicines related admission				✓		✓		
Social-related admission		✓		✓		✓	✓	
Capacity as defined by Mental Capacity Act		✓		✓		✓	✓	
Member of travelling community		✓				✓		
Elderly living alone							✓	
Decanting of medicines occurring				✓			✓	
Significant weight changes				✓		✓	✓	
Bariatric patient		✓				✓		
Nutritional status				✓				
Smoking status								Low strength of evidence
Transgender		✓		✓				
Housebound						✓		
Disability				✓		✓	✓	
Pain score				✓			✓	
Venous access patient / type of cannula								Cannulas routinely used for all study patients
Having received antibiotics in last 3 months						✓		
Frequency of GP contact	✓					✓		
Identifying if patient is on a risk register with general practitioner		✓				✓	✓	

Continued from previous page…

Predictor	Reason(s) for non-selection							
	High correlation with other predictor(s)	Occurs infrequently*	Not available when model intended to be used	Not clearly defined / subjective definition / not reproducible	Potential measurement error	Not part of standard datasets &/or not reliably recorded†	Related to problems encountered by patients in primary care	Other / comments
Pregnancy / breastfeeding		✓						
Requirement to manipulate the medicine before administration				✓				
Critical care admission		✓						
Staffing levels on ward / hospital				✓		✓		Related to 'process' of prescribing rather than patient specific
Communication problems across interfaces				✓		✓		
The days admitted / time of year								Related to 'process' of prescribing rather than patient specific
Number of patient transfers across wards						✓		

* Estimated occurrence <10% patients

† Not included in standard admission proforma at study sites and/or data available for <50% patients (based on a review of the patient records of 84 patients, 50 from Hospital A and 34 from Hospital B)

MRP = medication related problem

5.4 Discussion

Key findings

This research has involved joint consideration of evidence from previous research, expert opinion, and the methodological considerations for prognostic modelling to provide an evidence-based approach for selecting the candidate predictors for use in development of the MOAT. One hundred and eighteen potential predictors were considered, 59 from previous research and 59 suggested by survey respondents. The total number pre-selected for use in development of the MOAT was 18.

Comparison with previous literature

Of the 59 potential predictors identified by the literature review, 32 were not selected for the expert survey because they did not meet the methodological requirements for prognostic modelling. The most common reasons were not being clearly defined, standardised and reproducible, or not forming part of standard clinical datasets. These issues could be overcome by developing clear, objective definitions, then collecting the data specifically for study purposes, but this would impact on the usability of the MOAT in clinical practice, as it would require additional measurements to be made / recorded. There is the also the potential for alternative definitions to be used, leading to measurement error and reduced prognostic accuracy.

Of the 27 predictors included in the expert survey, a further 11 were excluded, three because they had low strength of evidence plus low rating for clinical relevance by survey respondents. Four were excluded as they were not clearly defined, standardised and reproducible, and/or not part of standard datasets. The remainder were excluded as they were correlated with other predictors, or occurred infrequently.

Eight of the 59 predictors suggested by expert survey respondents, such as ‘disability’ and ‘adherence with medicine taking’, had already been identified by the literature review, but the majority had not. Examples include ‘activities of daily living’, ‘patient health beliefs’, ‘nil by mouth / enteral tube use’, and ‘pain score’. There are three possible reasons for this.

1. The literature review included predictors for ‘adverse medication-related outcomes’, including adverse drug reactions (ADRs), adverse drug events (ADEs), medication related problems (MRPs), and medication errors (MEs). As the definition for MRPs is broader than the other outcome events it is likely that the predictors for MRPs are more varied.

2. Many of the suggested predictors did not meet the methodological requirements for predictors, often lacking a clear definition or standardisation. This may have resulted in them being excluded in previous studies.
3. Some predictors suggested by the survey respondents relate to the identification of patients who may require pharmacist input to prevent MRPs post-discharge, for example the 'Medication Regimen Complexity Index' which was designed to identify patients who would benefit from additional pharmaceutical input such as domiciliary reviews and special pharmacotherapy consultations[108]. This outcome was outside the scope of the literature review.

Interpretation and implications

This research highlights the value of using an expert survey to identify potential predictors as it permitted broader identification of potential predictors than literature review alone. It also provided expert advice on the need to include topical medicines in the analysis, which may otherwise have been omitted, and provided suggestions regarding how to categorise diagnoses and high-risk medicines.

It is of note that 32 of the 59 potential predictors identified by the literature review, and 57 of the 59 predictors suggested by the expert survey respondents were excluded from the final selection, either due to methodological limitations, and/or because they relate to the risk of MRPs post discharge, which is outside the scope of the present study. This raises two issues: (1) recognition that it is not possible to model the impact of all potential predictors; and (2) pharmacists may need to combine the MOAT with alternative methods of patient prioritisation to identify patients at risk of alternative outcome events.

Steyerberg advises that it is appropriate to omit predictors if their effect cannot be reliably estimated[70], similarly Sullivan *et al*[100] acknowledge the need to restrict predictors to those that are 'readily available in clinical practice and precisely measured'. Steyerberg recommends that if potentially highly significant predictors are excluded the model should be presented with an appropriate warning[70]. This will be an important consideration when introducing the MOAT into clinical practice.

A potential area for future research may be to investigate the impact of combining the MOAT with other triggers for pharmacist review that we are unable to model, such as patients with swallowing difficulties, or those receiving end of life care. Another potential area may be to investigate the benefits of combining the MOAT with tools

aimed at identifying patients at risk of MRPs post discharge, such as the PREVENT tool[109], to provide a more holistic approach to prioritisation.

Strengths and limitations

Strengths of the research presented in this chapter include the evidence-based approach, and high number of survey responses.

A potential limitation was the use of a simple count to identify predictors previously reported as prognostic. While this is a recognised method to identify predictors[70], it does not take into account differences in study design, study quality, or magnitude of effect. In addition, the categorisation of strength of evidence, based on the proportion of studies that found an association, does not reflect the number of studies investigating each predictor. For example, both gender and social deprivation are categorised as having low strength of evidence, but gender was included in 19 studies, and social deprivation in only one; this could further reduce the reliability of the assessment. I accounted for these potential weaknesses by not relying solely on the strength of evidence to select candidate predictors. Limitations for the expert survey include the use of convenience sampling, the 'infinite' target population which precludes calculation of a response rate, and the potential impact of volunteer bias.

5.5 Conclusion

This research resulted in the pre-selection of 18 candidate predictors for use in development of the MOAT. The method used theoretical knowledge, clinical understanding, and methodological considerations, with the aim of increasing the predictive performance[94], and clinical credibility[56] of the MOAT. In addition the research presented in this chapter highlighted that not all potential risk factors can be included in prognostic modelling, which will need to be considered when introducing the MOAT into clinical practice.

The next chapter will explore the data collection for model development in more detail, including the need to develop clear definitions for the selected candidate predictors.

Chapter 6: Data collection for model development

6.1 Introduction

As discussed in chapter 2, many published prognostic model studies have been criticised in terms of methodological shortcomings[55 56] and poor reporting[54 57], limiting their usefulness. In response, the PROGnosis RESearch Strategy (PROGRESS) group published recommendations to improve the quality and impact of prognosis research[53]. In relation to data collection, they recommended 'greater efforts to understand and improve the quality of clinically collected data', and a need for better reporting to 'improve transparency' and 'identify good-quality from low-quality research'[53]. The PROGRESS recommendations were followed by publication of the 'CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies' (CHARMS)[60], and the 'Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis' (TRIPOD) statement[57]. While neither prescribe how to develop a prognostic model, CHARMS highlights potential sources of bias, and TRIPOD summarises the qualities of good studies, and highlights inappropriate approaches that should be avoided. Both also provide guidance on the level of reporting required to permit quality to be adequately assessed.

The aim of the work presented in this chapter was therefore to describe the methodology, methods, and results of data collection, using approaches informed by the PROGRESS[53 55], TRIPOD[94], and CHARMS[60] recommendations. The objectives were to:

- enhance the potential clinical credibility and usability of the Medicines Optimisation Assessment Tool (MOAT^{TM}) by selecting appropriate definitions for the outcome measure and candidate predictors;
- select methods that minimise the risk of bias (selection and information biases);
- explain and justify participant inclusion and data collection methods;
- provide descriptive analysis of the sample population to permit potential users of the MOAT to assess applicability and generalisability;
- demonstrate the rigour and consistency employed throughout all stages of data collection.

This chapter includes both the methods and results for this part of the research.

6.2 Methodological considerations

When designing this study it was necessary to make a number of fundamental methodological decisions related to data collection. These included the choice of:

- outcome measure (the adverse medication-related outcome to be predicted by the MOAT);
- data collection method for the outcome measure;
- classification system to categorise the outcome data;
- definitions and/or categories for the candidate predictors selected in chapter 5 (including data collection methods where methodological decisions required).

Each of these is described below.

6.2.1 Selection of outcome measure

As discussed in chapter 3, various outcome measures have been used in studies to predict adverse medication-related outcomes including adverse drug reactions (ADRs), adverse drug events (ADEs), medication related problems (MRPs), and medication errors (MEs). Moons *et al*[94] advise that the choice of outcome event is a critical factor in determining the clinical relevance of a prognostic model, making this a key early consideration. The potential advantages of ADEs or ADRs over MRPs or MEs are as follows:

- ADRs and ADEs are objective measures of harm, whereas MRPs and MEs include potential harm, therefore are subjective;
- while ADRs and ADEs represent harmful events, only a small proportion of MRPs or MEs result in actual adverse medication-related outcomes[78 83]. It could therefore be argued that patient safety initiatives that target patients at risk of actual rather than potential harm may be more efficient;
- unlike ADRs and ADEs, MRPs and MEs are proxy measures for patient-oriented outcomes, which limits the ability to directly relate a reduction in MRPs or MEs to patient outcomes.

While these are valid arguments, the proposed purpose of the MOAT is to target patients most in need of pharmacists' input, and as pharmacists routinely identify and resolve potential in addition to actual adverse medication-related outcomes, a prediction tool that more closely reflects clinical practice is likely to have greater clinical credibility. Another consideration was the ability to collect valid outcome data given that regular patient review by ward-based clinical pharmacists is standard practice at the

proposed study sites, as it is in many United Kingdom (UK) hospitals. Pharmacist intervention therefore had the potential to prevent harmful events from occurring, so reducing ADE occurrence (i.e. outcome data would only include events that were not intercepted, or could not be prevented by pharmacist intervention). Although it would be possible to eliminate this source of bias by removing the clinical pharmacy service during the study, withholding standard care would raise significant ethical issues. It was therefore decided to look at MRPs or MEs.

Regarding the choice between MRPs and MEs, MRPs encompass a wider range of events, including any untoward medication related outcome irrespective of whether an error occurred. The use of MRPs as the outcome measure also permits targeting of patients with unrealised benefits (as discussed in chapter 2). I therefore selected MRPs as the outcome measure as I believe it will align the MOAT more closely with pharmaceutical care practice[110], so enhancing clinical credibility.

Severity and preventability of outcome measure

Previous research suggests that a significant proportion of hospitalised patients experience MRPs (for example Blix *et al* reported a prevalence of 81%[78]), many of which are of limited clinical significance. A model developed to predict MRPs would therefore lead to a high proportion of patients being labelled as high-risk, potentially leading to inefficient workload management. An option suggested in chapter 3 was to select a clinically significant outcome measure, based on severity. I therefore chose to severity rate all MRPs, and only use moderate or severe MRPs for model development (as described in chapter 8). Similarly, the aim of the MOAT is to target patients with MRPs amenable to pharmacist intervention (i.e. those with preventable MRPs). Unpreventable MRPs, such as ADRs that could not have been anticipated, were therefore excluded from MOAT development. As no established grading system for MRP preventability is available, I considered two possible methods: the criteria provided by Schumock and Thornton[111], and the 'P Method'[112]. I concluded that both methods were developed for ADRs, most of which are unpreventable, whereas the majority of MRPs are inherently preventable. Neither method was therefore appropriate for the present study. Pharmacists were therefore asked to review each MRP at the point of identification to assess whether it was preventable, expressed as a dichotomous variable of yes or no.

6.2.2 MRP data collection

Previous research into the detection of prescribing errors, a subset of MRPs, has shown that the observed prevalence is extremely dependent on the method of detection[113]. I chose to use prospective identification by pharmacy staff for this study because: (1) the purpose of the study was to develop a prognostic model for MRPs that can be identified during routine clinical practice; (2) it would permit the identification / inclusion of MRPs that are not routinely recorded in medical notes, such as potential prescribing or administration errors that are intercepted; and (3) it would permit the MRPs to be identified by staff personally involved in the care of the study patients, increasing clinical and practical relevance.

6.2.3 MRP classification

A number of classification systems for MRPs have been developed, which vary in the definition used, and method to classify causes and outcomes. As there is no universally accepted classification system, and perceived deficiencies with many published classification systems, I used the aggregated classification system recently developed by Basger *et al*[114] (Table 13). This is based on the most commonly used systems (Pharmaceutical Care Network Europe[20], Cipolle *et al*[115], Westerlund[116], DOCUMENT[117], Norwegian[118], and the individualised Medication Assessment and Planning / iMAP tool[119]), and provides a comprehensive classification system based on the causes of MRPs, thereby preventing any potential confusion between MRP causes and outcomes.

Table 13 – Basger's medication related problem classification system

1. Drug selection
1.1 Inappropriate drug
1.2 No indication for drug / duplication
1.3 Interaction (drug-drug, or drugs and food / alcohol)
1.4 Indication not treated / missing therapy
1.5 More cost effective drug available
1.6 Synergistic / preventive drug required and not given
2. Drug form
2.1 Inappropriate or suboptimal drug form
3. Dose selection
3.1 Drug dose too low
3.2 Drug dose too high
3.3 Dosage regimen not frequent enough
3.4 Dosage regimen too frequent

3.5 Dose needs adjustment to organ function or change in disease state
3.6 Dosage instructions unclear, incomplete or not understood by patient / carer
4. Treatment duration / withdrawal
4.1 Duration of treatment too short
4.2 Duration of treatment too long
5. Drug use process
5.1 Inappropriate timing of administration / dosing by prescriber; administration error by nurse
5.2 Drug underused / under-administered
5.3 Drug overused / over-administered
5.4 Drug not taken / administered at all
5.5 Wrong drug taken by patient
5.6 Drug abused
5.7 Patient or nurse uses drug incorrectly through lack of knowledge or barriers (e.g. swallowing, dexterity)
5.8 Adequate information not provided or not understood or misunderstood or not followed
5.9 Drugs stored inappropriately / expired drug administered / preparation error
6. Logistics
6.1 Prescribed drug not available
6.2 Drug order incorrect, incomplete, poorly legible / illegible / illegal / incorrect / allergy status incomplete
6.3 Error in drug selection
7. Monitoring
7.1 Monitoring too frequent
7.2 No or too infrequent monitoring
7.3 Inappropriate test ordered
7.4 Patient unable to attend / pay for monitoring
8. Unexpected reaction / adverse drug reaction (ADR) / no obvious cause
8.1 An ADR occurred
8.2 No obvious cause of treatment failure

Three of Basger's MRP subcategories were not used for the present study as they relate to primary care:

- 3.6 'dosage instructions unclear, incomplete or not understood by patient / carer';
- 5.8 'adequate information not provided or not understood or misunderstood or not followed';
- 7.4 'patient unable to attend / pay for monitoring'.

I also added a category for 'inappropriate abrupt withdrawal of a medicine', as I did not feel that this was adequately captured by Basger's classification system.

6.2.4 Selection of candidate predictor definitions / categories

Once the candidate predictors had been selected (as described in chapter 5), it was necessary to choose appropriate definitions for each of these. Unambiguous definitions are needed to improve the accuracy and replicability of predictions[60 98], but clinical credibility can be reduced if definitions are not clinically relevant, or too difficult to apply in practice[102]. My aim was, therefore, to: (1) select definitions commonly used in clinical practice; (2) use data that are readily available (to avoid the need for additional measurements, which could be perceived as inconvenient and/or costly[102]); and (3) avoid the need for complex calculations or categorisation (to ensure ease of use in clinical practice[102]).

It was also necessary to pre-select categories for the categorical predictors with more than two categories (primary diagnosis and high-risk medicines), to reduce the risks of model overfitting (associated with using too many variables[60]), and selection bias (caused by data-driven analysis[62]).

The selection of predictor definitions and categories is described below.

6.2.4.1 Laboratory results

Laboratory results vary throughout a patient's admission to hospital, it was therefore necessary to select a time point for data collection to ensure consistency and reproducibility.

Prognostic modelling investigates the relationship between future outcomes (endpoints) among people with a given baseline health status (startpoint)[120]. For the MOAT study the 'startpoint' was admission to a hospital medical ward, I therefore used the first documented laboratory results following admission. In six cases (all from Hospital A), no results were available during the hospital stay (i.e. serum sodium, potassium, creatinine, albumin, white cell count and liver function tests), but results were available from the preceding week. I consulted the MOAT project steering group (comprising PhD supervisors, clinical supervisors, patient / public representatives) to discuss whether to use these results for the study. It was agreed that this was appropriate, on the basis that the results were the only ones available to the healthcare team during admission. This represented 0.4% of patients (i.e. six from a total of 1,503).

Details of how laboratory results were used to estimate renal and liver function are given next.

Renal function

The two prediction formulas commonly used to estimate renal function are the Cockcroft-Gault, and modified Modification of Diet in Renal Disease (MDRD) equations[121]. There is relatively good correlation between the two for patients of average weight[122], but results are not interchangeable, particularly at extremes of body weight[121]. A key deciding factor in selecting which equation to use for this study was the availability of data. Both equations require serum creatinine, age and gender, but Cockcroft-Gault adjusts for body weight, whereas the modified MDRD adjusts for race. My original intention was to use Cockcroft-Gault, as this is more frequently used by pharmacists in clinical practice (to adjust medicine dosages in renal impariment[122]). However, in reviewing the study data I found that weight was unavailable for 186 (12.4%) of 1,503 patients. Race was more reliably recorded, with data unavailable for only 96 (6.4%) of patients.

A further consideration was the increasing use of the modified MDRD equation by laboratories throughout the UK to routinely report renal function whenever a measurement of serum creatinine is requested[123], as is the case at Hospital B.

I therefore chose to use the modified MDRD equation to estimate renal function for this study, calculated using the following formula:

Estimated GFR (ml/min/1.73m^2) = 186 x (creatinine / 88.4)$^{-1.154}$ x (age)$^{-0.203}$ x (0.742 if female) x (1.210 if black)

as recommended by National Institute for Health and Care Excellence[124], Renal Association (UK)[125], and used by Renal Drug Database[122].

I discussed how to handle the missing ethnicity data with Dr Li Wei (academic supervisor), and we decided that where ethnicity was not recorded, patients would be categorised as non-black for the purposes of estimating renal function. This decision was taken due to the relatively small number of admissions with missing ethnicity data, and the low number of ‘black’ patients within the remaining 1,407 admissions (45 patients, i.e. 3.2%). Although this proportion may not be representative of the patients with missing ethnicity data (as individual ethnic groups may be more or less likely to withhold ethnicity information), if one assumes that 3.2% of the admissions with missing ethnicity data were black, this would equate to three additional black patients (i.e. 3.2% of 96 admissions). Given the potentially small number, it was agreed that classifying all missing ethnicity data as ‘non-black’ would not significantly affect the overall predictor-outcome relationship.

Liver disease definition

Liver disease is not straightforward to quantify due to the variation in liver function tests (LFTs) dependent on the type and stage of disease[126]. I therefore chose to treat liver disease as a binary variable (yes/no). To establish a suitable definition, I reviewed the definitions used by studies included in the literature review (chapter 3), and pharmacy triage tools currently in use within hospitals in the UK: NHS Greater Glasgow and Clyde, Royal Cornwall Hospitals NHS Trust, and Leeds Teaching Hospitals NHS Trust. I became aware of these consensus derived tools during a meeting at NHS England (July 2016), held to discuss pharmacy prioritisation methods used in UK hospitals (as part of their work on developing seven day clinical pharmacy services in acute hosptials[36]).

Liver disease was included as a risk factor in the following triage tool / studies, but was not defined:

- Leeds Teaching Hospitals NHS Trust (triage tool);
- Hickson *et al*[51];
- Kaufmann *et al*[38];
- Onder *et al*[41];
- Evans *et al*[79].

The definitions used for liver disease by the remaining triage tools / studies are summarised in Table 14.

Table 14 – Definitions used for liver disease

Source	Liver disease definition
NHS Greater Glasgow and Clyde triage tool	Severe hepatic impairment defined as liver function tests (LFTs) ≥ 3 times upper limit of normal Moderate liver impairment defined as LFTs elevated from normal but < 3 times upper limit of normal
Royal Cornwall Hospitals NHS Trust triage tool	High priority (level 3) patients: LFTs: alkaline phosphatase (ALP), alanine aminotransferase (ALT), bilirubin ≥ 3 times upper limit of normal
Blix *et al.* (2004)[78]	LFTs – aspartate aminotransferase (AST) or ALT 3 times above normal value
Wilmer *et al.* (2015)[90]	Child-Pugh score
O'Connor *et al.* (2012)[88]	Liver disease defined as synthetic liver dysfunction, or liver injury with raised transaminases greater than twice the normal range, or documented liver disease
Beckett *et al.* (2012)[87]	Liver impairment taken as ALT or AST > 40mg/dL (standard reference range not given)

As shown in Table 14, a range of tests and/or cut-points have been used to define liver disease. Wilmer *et al*[90] used the Child-Pugh score, whereas the reminder used LFTs. I chose to use LFTs rather than the Child-Pugh score as this is consistent with clinical practice, as shown by the use in two of the pharmacy triage tools already in use in the UK (NHS Greater Glasgow and Clyde, and the Royal Cornwall Hospitals NHS Trust). In addition, the Child-Pugh score is a measure of prognosis of chronic liver disease, therefore not routinely assessed.

LFTs is a term used for a range of tests of liver function, including aspartate aminotransferase (AST), alkaline phosphatase (ALP), alanine aminotransferase (ALT), gamma-glutamyl transpeptidase (GGT), bilirubin, and serum albumin[126]. As shown in Table 14, the LFTs commonly used to establish a diagnosis of liver disease were AST, ALT, ALP and bilirubin. Blix *et al*[78] and Beckett *et al*[87] used AST, but this was not an option for the present study as it was not routinely measured at either of the study sites. I therefore chose to use ALP, ALT, and bilirubin, as this is consistent with clinical practice at both study sites, and the Royal Cornwall Hospitals NHS Trust. I chose to use a cut-point of greater than or equal to three times normal ranges. This is consistent with the triage tools used by NHS Greater Glasgow and Clyde, and the Royal Cornwall Hospitals NHS Trust. It was also the cut-point used by Blix *et al*[78].

Where no LFT results were available (i.e. where tests not performed) I assigned the outcome as ‘normal LFTs’ on the assumption that the medical team had not considered the tests necessary. This occurred for 26 (1.7%) of 1,503 patients. Where not all results were available (due to the laboratory being unable to analyse the sample for specific tests, for example where sample haemolysed) I also assigned the result as ‘normal LFTs’ if the remaining result(s) were within normal range. This occurred for 71 (4.7%) of the 1,503 patients.

Given the variation in liver function tests dependent on the type and stage of disease[126], and the lack of laboratory data for some patients (as detailed above), I also chose to include ‘documented liver disease’ as an alternative marker of liver impairment, as used by O’Connor *et al*[88] (see Table 14). For this, I reviewed the International Classification of Diseases (ICD) data for all study patients (diagnoses and comorbidities). Patients with one or more of the following ICD codes were assigned the outcome ‘documented liver disease’:

- acute and subacute hepatic failure;
- acute viral hepatitis;

- alcoholic cirrhosis of liver;
- alcoholic fatty liver;
- alcoholic fibrosis and sclerosis of liver;
- alcoholic hepatic failure;
- alcoholic liver disease, unspecified;
- autoimmune hepatitis;
- chronic passive congestion of liver;
- chronic viral hepatitis B;
- chronic viral hepatitis C;
- fatty change of liver;
- hepatic failure;
- hepatomegaly;
- hepatorenal syndrome;
- liver disease, unspecified;
- other and unspecified cirrhosis of liver;
- other specified diseases of liver;
- other specified inflammatory liver diseases;
- portal hypertension;
- malignant neoplasm: liver cell carcinoma;
- secondary malignant neoplasm of liver and intrahepatic bile duct;
- malignant neoplasm, liver cell carcinoma;
- abnormal results of liver function studies.

Study patients were therefore recorded as having liver disease if they had an abnormal LFT result, and/or documented liver disease.

6.2.4.2 Comorbidity

Comorbidity is a complex entity, with little consensus on the best approach to measure it[127]. Various methods have been developed including:

- organ-based approaches, such as the Modified Cumulative Illness Rating Scale (CIRS)[128];
- weighted indices, such as the Charlson Index[129];
- prevalence of individual conditions, such as the Elixhauser system[130];
- simple counts of individual conditions[127].

After reviewing the above options I chose to use a simple comorbidity count. The reasons for this were:

- the Charlson Index and Elixhauser system were developed to predict the risk of mortality, and recent studies suggest they are poor predictors of nonfatal outcomes[47]. They are also based on a limited range of conditions, rather than being comprehensive measures of disease burden. For example, the Charlson Index excludes arrhythmias and thyroid diseases, and Elixhauser excludes cerebrovascular disease and previous myocardial infarction (Appendix A6.1 shows a detailed breakdown of the comorbidities included in CIRS, Charlson, and Elixhauser);
- the Charlson Index and CIRS require knowledge of disease severity. As comorbidity data for the MOAT study were obtained from ICD coding, it was not possible to perform this level of assessment. In addition, neither score is routinely used in clinical practice, potentially reducing ease of use of the MOAT;
- the Elixhauser system uses 30 dichotomous variables, while this would provide study-specific regression coefficients for each comorbidity, it could jeopardise regression modelling due to overfitting[70]. In addition, Elixhauser uses ICD coding from administrative systems. This is not available until after hospital discharge, so would not be available to pharmacists when the MOAT is intended to be used (unless patients had previous hospital admissions);
- a simple comorbidity count uses data that are routinely collected in clinical practice, are available at the point of hospital admission, and do not require complex manipulation. In addition, Steyerberg advises they 'may be rather robust and generalise well to new patients'[70].

Having decided to use a simple comorbidity count, I then reviewed the available ICD comorbidity data to assess usability, and identified the following issues:

- some ICD codes are not specific to conditions (e.g. 'abnormal finding of blood chemistry', 'malaise and fatigue');
- some are not directly related to illness (e.g. 'inadequate housing', 'allergy to penicillin');
- duplication can occur (e.g. 'atrial fibrillation and atrial flutter, unspecified' and 'atrial fibrillation and atrial flutter' used for the same patient);
- comorbidities related to the same underlying condition can be coded separately for the same patient (e.g. 'chronic obstruct pulmonary disease', and 'emphysema');

- temporary conditions are included as comorbidities (e.g. ‘pregnant state’, ‘pneumonia’), or relatively minor conditions (e.g. ‘conjunctivitis’, ‘otitis externa’).

This is likely to relate to ICD codes being used for a wide variety of signs, symptoms, abnormal findings, complaints and social circumstances[131]. It was therefore necessary to modify the ICD data to reduce the risk of double counting, and inclusion of conditions that may not be considered to be true comorbidities in clinical practice. I therefore reviewed the medical records of 96 randomly selected patients from Hospital A, (approximately 10% of study patients from that site), to compare ICD coding with the comorbidities listed on each patient’s discharge prescription (taking the latter as a source consistent with clinical practice). I found that the two sources of data were not directly comparable. As expected, the number of comorbidities was higher based on the ICD data (mean 8.2 comorbidities per patient compared to 4.3 for data from discharge prescriptions), for the reasons discussed above. As a result I chose to exclude the following ICD codes from the study comorbidity count:

- temporary conditions (e.g. infections);
- non-specific symptoms (e.g. cough);
- abnormal laboratory or other clinical findings (e.g. abnormal glucose tolerance test);
- procedures performed (e.g. cardiac catheterisation);
- trauma (e.g. unspecified injury of forearm);
- lifestyle factors (e.g. smoking, alcohol or illicit drug use);
- findings unrelated to illness burden (e.g. allergy status).

I also restricted the comorbidity count to conditions present prior to the current admission, that is, I excluded the primary diagnosis, as this may not be established at the point of admission to hospital. This is consistent with the definition used by Elixhauser[130]. I also grouped similar comorbidities to prevent double counting. I used the 14 CIRS organ-based categories as a framework, as this provided a comprehensive list of conditions associated with chronic illness burden[128], then sub-divided each category to permit more than one condition within an organ-based group to be counted individually. I created an additional category for tumours / malignancies; CIRS includes these within each organ-based system, but I grouped them to simplify the count, and prevent double counting of primary and secondary malignancies. This created 76 separate categories (shown in Appendix A6.2). These categories were used to recode the comorbidity related ICD data for all study patients, then duplicates removed to give a final comorbidity count for each patient.

6.2.4.3 Dementia

In addition to the inclusion of dementia in the comorbidity count, it was also selected as a separate candidate predictor. ICD coding was used to establish a documented diagnosis of dementia. All study patients with one or more of the following ICD codes were classed as having a history of dementia, expressed as a dichotomous variable of yes or no:

- dementia in Alzheimer disease;
- vascular dementia;
- unspecified dementia;
- delirium superimposed on dementia;
- Alzheimer disease with early onset;
- Alzheimer disease, unspecified;
- circumscribed brain atrophy;
- dementia in Alzheimer disease with early onset;
- dementia in Alzheimer disease, atypical or mixed type;
- dementia in Alzheimer disease, unspecified;
- dementia in other specified diseases classified elsewhere;
- dementia in Parkinson disease;
- dementia in Pick disease;
- other Alzheimer disease.

I excluded 'delirium', 'mild cognitive disorder', and 'signs involving cognitive function', due to their non-specific, and potentially temporary nature. I also excluded ‘senility’, due to its potentially subjective definition (covering both physical and mental age-related decline)[132].

6.2.4.4 Use of medicines

Three medicine related predictors were selected for development of the MOAT; the number of medicines, route of administration, and type of medicine prescribed (shown in Table 11, chapter 5). Before selecting definitions for these, it was necessary to define ‘medicine’.

During hospitalisation, medicines can be prescribed to be given in three ways, ‘STAT’ (abbreviated from the Latin word ‘*statum*’, meaning to be given immediately[132]), ‘PRN’ (from the Latin phrase ‘*pro re nata*’, meaning as required[132]), or as a regular, on-going

prescription. I chose to restrict the study to medicines prescribed to be given on a regular basis for the following reasons:

- STAT medicines may pose different risks to regular medication due to the range prescribed, e.g. vaccines and medicines used for diagnostic purposes;
- PRN medicines may be prescribed, but administered infrequently or not at all, which could falsely inflate the medicine count. For example, medicines such as analgesics and antiemetics are often prescribed to be given only if symptoms occur.

I also excluded the following:

- dietary products and emollients, on the basis these are non-medicated;
- wound dressings, as although sometimes medicated, they are not used as a means of medicines administration;
- oxygen therapy, and blood products, as these are not recorded as part of the medication records at all hospitals;
- 'water' or 'saline flushes'.

All other medicines prescribed to be administered on a regular on-going basis, including analgesics and antiemetics, were included, irrespective of the route of administration.

Conventions were used to ensure consistency in data recording (summarised in Appendix A6.3).

Number of medicines

As with laboratory results, the number of medicines prescribed will vary throughout a patient's admission to hospital, it was therefore necessary to select a time point to perform the medicine count. I chose to use the first full day of admission to hospital, as this provided 'startpoint' data, and included all medication prescribed to be given over a 24 hour period.

In some cases medicines were changed to a clinical equivalent on the first full day of admission, for example from intravenous to oral, or to a direct clinical alternative. Where this was clearly the case (i.e. where the same medicine was prescribed, but the route changed, or where it would not be clinically appropriate to use both medicines concurrently), this was counted as 'one' medicine.

Other conventions that were used were:

- combination medicines (e.g. solifenacin 6mg with tamsulosin 0.4mg modified release tablets) were counted as 'one' medicine;
- where more than one formulation of the same medicine was prescribed by same route (e.g. standard and modified release oral preparations, or suppository and enema), this was counted as 'one' medicine. If the same medicine was prescribed via different routes (e.g. mesalazine orally and rectally), both were counted separately;
- at Hospital A, medicines administered using a 24-hour subcutaneous syringe driver were prescribed as 'syringe driver' on the electronic prescribing record, irrespective of the number of medicines within the syringe. Syringe drivers were therefore counted as 'one' medicine at both sites;
- the count included all medicines prescribed to be given, even if withheld on the day the count was performed (e.g. once weekly medication not due to be administered on the day of the count, or medicines withheld for clinical reasons such as dehydration, or while awaiting results of other investigations).

Route of administration

Previous research suggests that the route of medicine administration is associated with adverse medication-related outcomes, with the strongest associations found with parenteral use[79 87 91]. This is supported by the National Patient Safety Agency alert, 'promoting safer use of injectable medicines'[133]. I therefore chose to assess the impact of parenteral medicine use, that is, administration via the intravenous, intramuscular, or subcutaneous route.

The following were excluded from this assessment:

- medicines prescribed as 'parenteral *or* oral', as it was not possible to establish whether the parenteral route had been used;
- prophylactic low molecular weight heparins, as they are routinely used in the majority of patients admitted to hospital medical wards[134];
- parenteral fluid replacement therapy.

For simplicity, parenteral medicine use was expressed as a dichotomous variable of yes or no.

Type of medicine

The risk associated with different medicines has been assessed in various ways in previous studies, including grouping by Anatomical Therapeutic Chemical (ATC)

codes[52 85], Institute for Safe Medication Practices (ISMP) high-alert medication[87], those with a 'narrow therapeutic index'[38 90], and by assessing individual medicines and/or classes[76 80].

Grouping medicines, rather than assessing individually, has the advantage of increasing the generalisability of results (as discussed in section 3.4), but the use of broad categories, for example ATC codes (where medicines are divided into different groups according to the organ or system on which they act[135]), can result in medicines with different pharmacological and chemical properties being grouped together. I therefore chose to focus on high-risk medicines / groups (discussed further in section 6.2.4.6).

6.2.4.5 Allergies

Data on all medicine-related allergies were collected for study patients. I considered using either a count of allergies, or treating it as a binary variable (yes/no), and chose the latter because: (1) a count may be dependent on breadth of medicine exposure, therefore correlated with comorbidity and/or age; and (2) this is consistent with previous research[38 41 42 44]. All non-drug allergies (e.g. food, latex) were excluded.

6.2.4.6 Categorisation of primary diagnosis and high-risk medicines

Primary diagnosis

My original intention was to categorise primary diagnosis using the ICD categorisation system, as used by other researchers[41 42], but after further consideration I chose to use an organ-based approach. ICD coding for primary diagnosis classifies diseases into 22 mutually exclusive chapters[131] based on 'the main condition treated or investigated during the relevant episode of healthcare'[136]. As ICD codes are allocated following the episode of care, codes for hospital inpatients are informed by investigations performed during admission. For example, up to 30% of patients with symptoms indicative of an acute ischaemic stroke may in fact have a 'stroke mimic', a non-vascular condition that also presents with acute neurological deficit, including seizures, brain tumours and infections[137]. As a result, further investigations, such as brain imaging, are required before the primary diagnosis can be confirmed. In this example, a confirmed stroke would be coded under the ICD chapter 'circulatory system', seizures under 'nervous system', brain tumours under 'neoplasms', with infections coded as either 'systemic infections' (under the chapter covering 'certain infectious diseases') or as a disease of the nervous system, depending on the site and type of infection. While ICD based categorisation may be useful for a retrospective analysis, the purpose of the MOAT is

to predict outcome events based on startpoint data (data available at the point of admission[120]). Use of an organ-based approach therefore permits categorisation ahead of definitive investigations, as all potential diagnoses are grouped together. For example, in the scenario above stroke and stroke mimics would be grouped as disorders of the nervous system. This has the potential to simplify use of the MOAT, and reduce the risk of misclassification.

I chose to use the CIRS organ-based classification system[128] as a framework to categorise primary diagnosis, as this is a recognised and comprehensive system, and is consistent with the approach used for the comorbidity count.

CIRS has 14 categories (shown in Table 15), but I chose to combine some categories (for example upper and lower gastrointestinal and liver diseases) to both simplify use of the MOAT in clinical practice, and reduce the risk of model overfitting (associated with using too many variables[60]). Remaining diagnoses were combined into an 'other' category on the basis they did not fall into the selected organ-based categories, or were symptoms or findings that are not specific to a diagnosis. This created eight categories, as shown in Table 15. Details of the specific diagnoses included in each category (based on the MOAT patients) are given in Appendix A6.4.

As discussed in section 5.3.2, the expert survey respondents suggested four diagnoses as potential risk factors for MRPs (diabetes, Parkinson's disease, immunodeficiency, and sepsis). Although it was not possible to include each of these individually, all fall within the primary diagnosis categories selected for the study.

Table 15 – Grouping used for primary diagnosis

Medicines Optimisation Assessment Tool category	**Modified Cumulative Illness Rating Scale categories**[128]
1. Cardiovascular system	Cardiac
	Hypertension
	Vascular-haematopoietic
2. Respiratory system	Respiratory
3. Gastrointestinal system	Upper gastrointestinal
	Lower gastrointestinal
	Hepatic
4. Genitourinary system	Renal
	Other genitourinary
5. Musculoskeletal-intergumentary systems	Musculoskeletal-intergumentary
6. Endocrine-metabolic diseases	Endocrine-metabolic
7. Nervous system and mental disorders	Neurological
	Psychiatric / behavioural
8. Other	Includes CIRS category 'eyes, ear, nose & throat' and symptoms or findings not specific to a diagnosis

High-risk medicines

The categorisation of high-risk medicines served two purposes; reducing the risk of model overfitting, and increasing the potential generalisability of the MOAT (as discussed in section 3.4).

I used the following sources to select the high-risk medicines categories for the MOAT study:

- two systematic reviews (a review by Suggett *et al*[64] that was specific for risk factors associated with the need for pharmaceutical intervention due to ADEs, ADRs and MRPs in adult hospitalised patients, and a review by Saedder *et al*[138] that identified high-risk medicines associated with MEs in adults, but was not specific to hospitalised patients);
- a meta-analysis by Boeker et al[139] of four studies of preventable ADEs in adult inpatients;
- a consensus study by Thomas *et al*[140] to identify prescribing indicators associated with potential harm in hospital settings;
- the consensus studies included in the literature review[38 46 48-51] (discussed in chapter 3);
- ISMP high-alert medication list[141];
- personal correspondence with staff at UK hospitals currently using pharmacy triage tools (NHS Greater Glasgow and Clyde, Royal Cornwall Hospitals NHS Trust, and Leeds Teaching Hospitals NHS Trust);
- suggestions received from the expert survey respondents (listed in Table 10, section 5.3.2).

Direct comparison between sources was difficult due to differences in the outcome measures, setting (hospital and primary care), the way medicines were grouped, and county specific issues (e.g. the ISMP high-alert medication list contains a number of medicines not routinely used in the UK). I therefore chose to include medicines as high-risk if there was evidence from more than one of the above sources, including a UK source. Five groups of medicines met these criteria, but were excluded for the following reasons:

- benzodiazepines / sedatives – excluded as often prescribed on an ‘as required’ basis during hospitalisation (whereas only medicines prescribed to be given on a ‘regular’ basis were included in the present study);

- non-steroidal anti-inflammatory drugs (NSAIDs) – also often prescribed ‘as required’ during hospitalisation;
- diuretics – not selected as high-risk in any of the consensus prediction tool studies, or pharmacy triage tools;
- anti-thrombotics – not selected as high-risk in any of the consensus studies, or pharmacy triage tools;
- corticosteroids – not identified as high-risk in any of the pharmacy triage tools, plus dosage and duration varies dependent on indication, which is likely to impact on the level of risk.

Clozapine, anti-retrovirals, and medicines for Parkinson’s disease were identified as high-risk, but I chose to group these, as I anticipated usage would be too infrequent to model each individually.

This resulted in 15 categories, as summarised in Table 16, section 6.3.4. Each was treated as a binary variable of yes or no. My original intention, as detailed in the published MOAT protocol[142], was to include ‘antibiotics’ as one of the 15 categories. I subsequently decided to change this to ‘antimicrobials’, as this permitted the inclusion of agents that act against all microbial organisms (i.e. antibiotics, antivirals antifungals, and antiprotozoal agents).

6.3 Methods

The methods are presented according to the TRIPOD reporting guidelines[57].

This section includes the source of data, selection of study participants, method of data collection for the outcome measure and candidate predictors, analysis of missing data, data entry checks for candidate predictors, statistical analysis, and ethical considerations. Each of these is described in turn below; the results are described in section 6.4.

6.3.1 Source of data

This prospective cohort study included patients admitted to two UK hospitals in South East England. The two study sites, Hospitals A and B, were chosen to increase generalisability of the MOAT as they have markedly different patient demographics.

Eligible patients were consecutively included at Hospital A from 28 April 2016 to 31 May 2016, and Hospital B from 19 October 2016 to 1 November 2016. All patients were followed up until discharge from hospital, or the date the study closed (two weeks after inclusion of the final study patient) whichever occurred sooner. A study close date was used to facilitate practicality in terms of data collection, while permitting data to be collected from admission to discharge for the majority of study patients (as the mean length of stay at the study sites was estimated at approximately six days).

6.3.2 Participants

The choice of participants was guided by the intended use of the MOAT, which is to identify adult patients at highest risk of MRPs during admission to a UK medical ward, irrespective of age. I therefore included all adults (aged 17 years and over). They were selected by means of being consecutive admissions to the medical wards (general, acute, and elderly medicine) at the study sites. At Hospital A there were 11 study wards (six general, one acute, and four elderly medicine). Hospital B had 19 study wards (six general, four acute, and nine elderly medicine). Patients admitted to other specialities such as surgery, maternity and paediatrics were excluded due to potential differences in the prevalence / type of MRPs in these patient groups. Patients admitted more than once during the data collection period were eligible to re-enter the study.

Patients were excluded if:

- their admission was for investigation only (as changes to medication would be minimal);

- they were not prescribed any medication during the admission;
- their entire admission was outside of core pharmacy working hours (i.e. 9am-5pm Monday-Friday) as these patients would be unlikely to receive review by a clinical pharmacist;
- their prescribing records were not reviewed by a clinical pharmacist during the admission (e.g. a patient who was present on a study ward during core pharmacy working hours but discharged before a clinical pharmacist was able to review their medication).

Patients admitted for investigation only, and those not prescribed medication during the admission, were excluded on the basis that they did not represent the target population for the MOAT. Those with admissions outside core pharmacy working hours, and whose prescribing records were not reviewed by a clinical pharmacist during their admission, were excluded on the basis that it was not possible to ascertain whether they experienced an MRP (using the chosen data collection method); inclusion of these patients may therefore have distorted the predictor-outcome relationship.

Patients were also excluded if their prescribing records and/or medical notes were unavailable. This was to ensure complete data were available for the medicine-related candidate predictors. The potential impact of excluding these patients is discussed in section 6.4.3 as part of the analysis of missing data.

The sample size target was 1,500 study admissions (1,000 from Hospital A and 500 from Hospital B). Section 9.2.3 in chapter 9 describes how this sample size was derived (as part of the methods used for model development). An additional 10% were included at each site to allow for potential losses due to exclusions.

The characteristics of the sample population are reported in section 6.4.2 to provide information on the context, case mix, and setting of the study[94]. This includes demographic data and the distribution of candidate predictors. Results are presented as percentage, mean or medians dependent on the type and distribution of data. Standard deviations or interquartile ranges (IQRs) are reported to establish variability, and ranges given for all continuous variables to inform values that will be compatible with the MOAT[94]. Due to the observational nature of the study, and retrospective collection of candidate predictor data, it was not possible to review measurement reliability, but reliability was assumed on the basis that data were collected for routine clinical purposes by healthcare professionals.

6.3.3 Outcome measure

As discussed in section 6.2.1, the outcome measure of interest for this study was MRPs. The definition used was 'all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome'[12 16 19 20]. Data were collected on all MRPs that occurred during the hospital stay, and then each was accessed for preventability and severity to identify the outcome event selected for MOAT development, namely moderate or severe preventable MRPs (MSP MRPs).

All pharmacists involved in the study were provided with training prior to commencement of the study (32 pharmacists at Hospital A and 44 at Hospital B). As MRP data were collected during daily ward visits and by staff in the centralised pharmacy dispensaries, the training was provided for all pharmacists who may have been required to clinically screen study patients' medication charts (i.e. ward-based pharmacists working routinely on the study wards, and all pharmacists who were involved in dispensary duties during the study period). The training comprised a 45-minute face-to-face session (delivered by myself) covering the purpose and design of the study, method of MRP data collection, discussion about the types of adverse medication-related outcomes covered by the MRP definition[12 16 19 20], and instructions on how to use Basger's MRP classification system[114].

Following this training, MRP data were identified and recorded by pharmacy staff at the study sites as part of their routine daily clinical assessment of patients, using a data collection form designed for this purpose (Appendix A6.5). Data were collected during daily ward visits (Monday to Friday 9am-5pm), and by staff in the centralised pharmacy dispensaries (Monday to Friday 9am-6.30pm and Saturday and Sunday 10am-4pm). The majority of clinical screening of medication orders occurred at ward level at both study sites, but data were collected in the centralised dispensaries to permit recording of MRPs identified outside routine ward pharmacy visits, for example medication requests made when the ward pharmacy team were unavailable. Pharmacy staff recorded data on all MRPs identified personally or through discussion with other healthcare professionals. The hospital incident reporting systems were also reviewed to check for any additional significant MRPs that were not identified by pharmacy staff.

In prognostic research it is recommended that the outcome event is assessed while blinded to the candidate predictors to prevent bias[60 93 94]. In the present study it was not possible to blind pharmacy staff collecting the outcome data to the patient's clinical

information (such as age, diagnosis and laboratory results) as this information formed part of their clinical assessment of patients. Despite this, pharmacy staff did not know which factors would be used as predictors, minimising the potential for this to influence their outcome assessment.

At the point of identification, pharmacy staff classified each MRP using Basger's classification system[114], and recorded whether they considered the MRP was preventable, expressed as a dichotomous variable of yes or no. I then performed a second check to increase consistency. To prevent judgement drift 'case law documents'[143] were developed and referred to at each stage.

Additional data were collected on the following:

- date MRP occurred and date resolved (to enable identification of duplicate reports);
- whether MRP was identified during ward visit or in the pharmacy department (to provide data on working practices at the study sites);
- whether MRP was resolved by pharmacy staff or other healthcare professionals (to differentiate between MRPs resolved by pharmacy staff and those identified and resolved by other healthcare professionals then reported to pharmacy staff retrospectively);
- stage in 'patient stay' when MRP identified, classified as during/before first ward review by pharmacist, during the remainder of the inpatient stay, or during clinical screening at discharge (to provide additional data on MRP occurrence to inform MOAT implementation);
- whether MRP was a medicines reconciliation discrepancy (as evidence suggests that patients are at increased risk of medication-related harm during transitions of care[3 7]). Medicines reconciliation discrepancies were defined as errors identified after creating the most accurate list possible of all medications a patient takes, including drug name, dosage, frequency, and route, and comparing that list against the physician's admission, transfer, and/or discharge orders, with the goal of providing correct medications to the patient at all transition points within the hospital[144].

A potential limitation was the possibility of incomplete data due to pharmacy staff being required to complete this work in addition to other routine duties. To minimise this, I worked closely with the study sites to ensure that data collection occurred at an optimal time in terms of staffing levels and workload. Staff involved in MRP data collection were also provided with initial training (as described above) to improve the consistency and

reliability of data collection. I reviewed all data collection forms daily and sought clarification where needed, and provided the pharmacy staff with on-going fidelity training.

I also recognised that identification of MRPs may vary depending on the knowledge, experience and skills of the pharmacists collecting data. To quantify this potential variability, a simulated 'MRP identification assessment exercise' was developed and used in a training scenario, discussed further in chapter 7.

Following anonymisation to maintain patient confidentiality and blinding, each potential MRP was independently assessed by an expert panel (consisting of a hospital pharmacist, senior nurse, consultant physician plus myself). The panel reached agreement by consensus on whether it was a true MRP (expressed as a dichotomous variable of yes or no). Where there was not full agreement, a final decision was made following discussion between panel members. To prevent 'judgement drift' a 'case law document'[143] was used as above.

6.3.4 Candidate predictors

The work undertaken in chapter 5 identified 18 candidate predictors for use in development of the MOAT.

All data on candidate predictors were collected retrospectively. Data were obtained from the information department at the study hospitals where possible, including demographic, diagnostic and comorbidity data. Laboratory data were extracted manually from the electronic reporting system used at both hospitals, Sunquest Integrated Clinical Environment (ICE)[145]. The remaining data were extracted manually from patient medical records. Hospital A has electronic medical and prescribing records; Hospital B has paper-based systems. Manual extraction of laboratory data were performed by a single data analyst at each study site, independently of the research team. Data from the patient medical records were collected by the independent data analyst at Hospital A, but due to the use of paper-based systems at Hospital B, and the need to read hand-written prescribing records, I extracted these data at Hospital B. All manually extracted data were entered directly into an electronic database. All data were recorded as reported, with no categorisation of continuous data.

As discussed in section 6.2.4.1, where no laboratory results were available during the hospital stay, results from the preceding week were used (if available). Similarly, if no

height or weight data were available (for the study admission) data were extracted for previous or subsequent admissions. If available, data on height were used irrespective of the date recorded (on the basis that height is relatively stable). Weight data were only used if results were available within one month of the study admission. These decisions were made following discussion with the MOAT project steering group.

Table 16 shows the candidate predictors that were pre-selected for development of the MOAT, with definition / measurement methods used (pre-selected to minimise heterogeneity and bias[94], as described in section 6.2.4).

Table 16 – Pre-selected candidate predictors

Variable	Details / categories
Age	Age at admission to hospital (in years)
Socioeconomic status	Based on the English indices of deprivation 2015 (Index of Multiple Deprivation Rank)
Previous allergy / adverse drug reaction	Binary (YES/NO)
Body mass index	First documented result following admission
Number of hospital admissions	Number of admissions to the study hospital in the previous 6 months
Primary diagnosis	From hospital clinical coding data (ICD-10 codes). Grouped into: 1. Cardiovascular system 2. Respiratory system 3. Gastrointestinal system 4. Genitourinary system 5. Musculoskeletal-intergumentary systems 6. Endocrine-metabolic diseases 7. Nervous system and mental disorders 8. Other (all other diagnoses combined)
Number of comorbidities	From hospital clinical coding data (ICD-10 codes)
History of dementia	Binary (YES/NO) From hospital clinical coding data (ICD-10 codes)
Number of medicines prescribed	Number of 'regular' medicines prescribed to be given on the first full day of admission to hospital. i.e. excluding 'when required' and 'once only' medicines, dietary products, non-medicated topical products (e.g. emollients), wound dressings
Parenteral administration route	Binary (YES/NO) Administration of one or more regular medicines via the parenteral route (intravenous, intramuscular, subcutaneous) during the hospital stay (excluding prophylactic low molecular weight heparins, fluid replacement therapy)
Use of 'high-risk medicines'	Binary (YES/NO) Prescribed to be given as a 'regular' medicine during the hospital stay: 1. Anticoagulants / direct oral anticoagulants 2. Therapeutic heparin 3. Anti-diabetic medication 4. Opiates (excluding codeine, tramadol, meptazinol & dihydrocodeine) 5. Systemic aminoglycosides & glycopeptides 6. Systemic antimicrobials (excluding aminoglycosides & glycopeptides) 7. Theophylline & aminophylline 8. Epilepsy medicines 9. Antipsychotics (excluding clozapine) 10. Immunosuppressants (excluding corticosteroids) 11. Cytotoxics 12. Lithium 13. Antiarrhythmics 14. Antidepressants 15. Other (clozapine, anti-retrovirals, medicines for Parkinson's disease)

Continued from previous page…

Variable	Details / categories
Renal function	Glomerular filtration rate calculated using the modified Modification of Diet in Renal Disease equation (using first documented results following admission)
Liver disease	Binary (YES/NO) Liver disease defined as ALT / ALP and/or bilirubin ≥ 3 times normal range and/or documented liver disease Laboratory results were the first documented results following admission Documented liver disease was established from hospital clinical coding data (ICD-10 codes)
Serum albumin	First documented result following admission
Serum potassium	First documented result following admission
Serum sodium	First documented result following admission
White cell count	First documented result following admission
Platelet count	First documented result following admission

ICD = International Statistical Classification of Disease, ALT = alanine aminotransferase, ALP = alkaline phosphatase

6.3.5 Analysis of missing data

To assess the potential impact of missing data I calculated the number of admissions with missing data, number of values missing, and number of missing values for each variable (i.e. each set of data collected). I then compared characteristics for admissions with missing values and those with completely observed data to inform possible reasons for the missingness[94].

6.3.6 Candidate predictor data entry checks

In prognostic research it is recommended that data on candidate predictors is collected blind in terms of knowledge of the outcome event and other predictors[60 94]. This is particularly important when subjective judgement is required as it prevents the assessment being influenced, which could artificially increase the associations between the predictors and outcome events. Full blinding was not possible for this study as both myself and the independent data analysts were not blinded to all other predictor data, and I was not blinded to the MRP status. It was anticipated that this would have minimal impact on the accuracy of data collection as all candidate predictors selected for this study were objective measurements that are independent of observer interpretation; subjective judgement was therefore not required. In addition all candidate predictors were recorded contemporaneously during the admission as part of routine care / documentation, therefore without knowledge of the MRP status. To identify any possible bias, and to assure the accuracy of data collection, I performed a

double check of data entry. This involved an on-going check of a randomly selected 10% sample of study patients (selected using a random number generator). Sixteen data items were checked for each patient. This included four from prescribing records, nine from laboratory reports, and three from patient medical records:

- prescribing records – medicines reconciliation completed (yes/no), list of allergies, list of medicines prescribed to be given on first full day of admission, and additional medicines prescribed during admission;
- laboratory reports – serum sodium, potassium, creatinine, albumin, bilirubin, ALT, ALP, white cell count, and platelet count;
- patient medical records - weight, height, and body mass index.

As the data entry checks were performed at regular intervals throughout data collection it was possible to refine data entry where necessary. Accuracy was calculated as the percentage of data items recorded correctly.

6.3.7 Statistical analysis of data

Statistical analyses were performed to test for differences between study sites in the characteristics of study admissions, MRP identification processes (such as whether identified during ward duties or in the pharmacy department), the proportion of each MRP subcategory, and the prevalence of MRPs; this was to inform the potential generalisability of the MOAT. The characteristics of admissions with and without missing data were also compared to identify the possible missingness mechanism, so informing how missing data should be handled during MOAT development.

The Mann-Whitney U test and two-sample t-test were used for numeric data, selected based on whether data distribution was compatible with normality. Chi-square or Fisher's exact tests were used to compare proportions, with the choice based on the *expected* distribution of results in the absence of association, with Fisher's exact test used when an expected cell frequency was less than five. Where appropriate, the Bonferroni correction was applied to the probability (*p*) values to account for the risk of type I (false positive) errors associated with multiple analyses[146]; Bonferroni corrected *p* values were calculated based on the number of comparisons (to maintain the critical *p* level over all tests at 0.05).

6.3.8 Ethical considerations

This study was approved by the Proportionate Review Service Sub-Committee of the NHS Research Ethics Committee Wales REC 7 (16/WA/0016), and Health Research

Authority (HRA) (project ID 197298). Informed written consent was not required as I held a contract of employment with both study sites.

All relevant policies and guidance related to confidentiality were followed to ensure the confidentiality of all patient-identifiable or confidential information. This included the use of study codes to allow data to be pseudonymised, recording the minimum information necessary, only sharing patient identifiable data with healthcare staff directly involved in the care of the patient, storing data securely, and ensuring that patient identifiable data was not disclosed in publications / presentations.

6.4 Results

The results are presented according to the TRIPOD reporting guidelines for prognostic model studies[57], and STROBE guidelines for the reporting of observational studies[147]. Recommendations for reporting the analysis of missing data have also been followed[148].

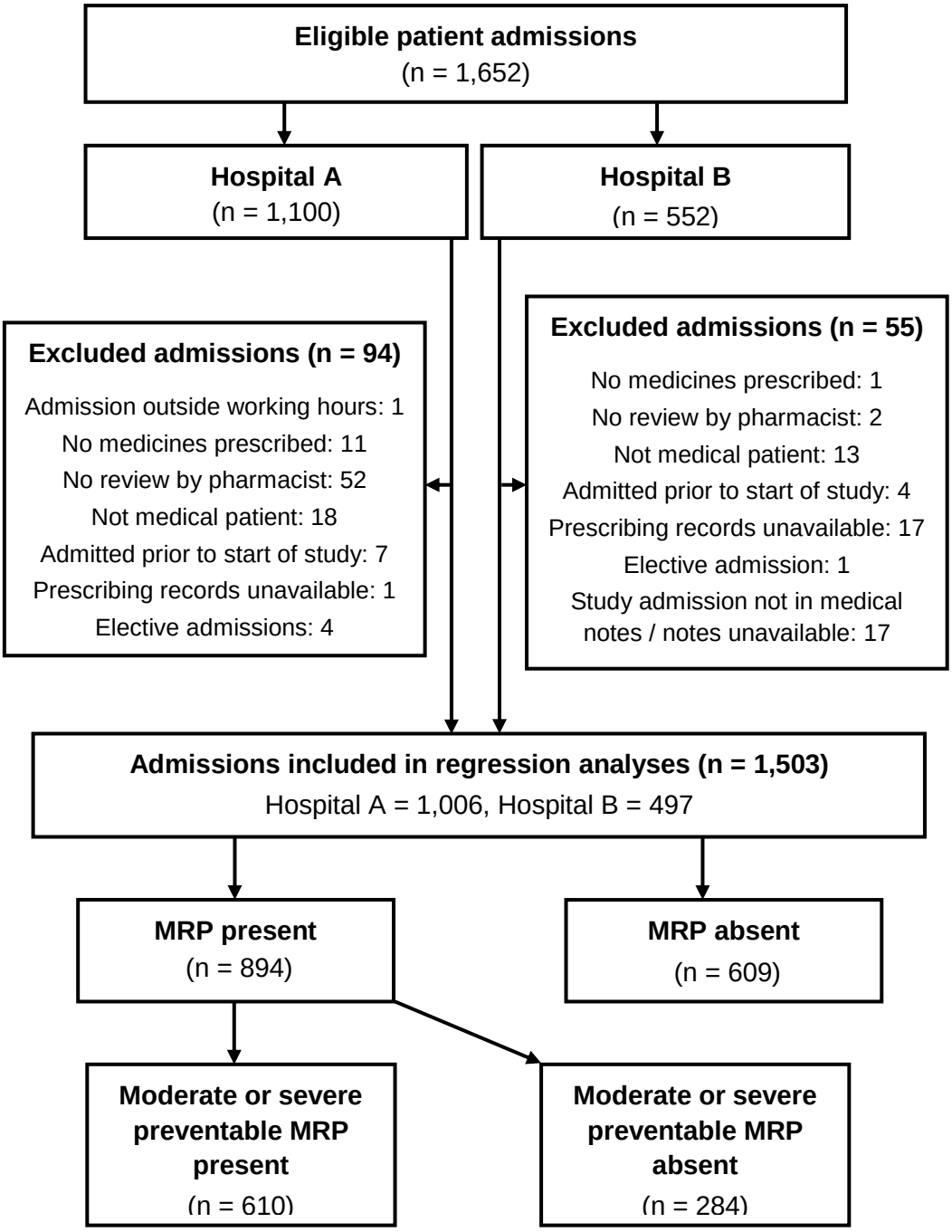
This section includes the flow of patients through the study, key characteristics of study admissions, analysis of missing data, results of data entry checks, and descriptive analysis of MRP data. The results of the preventability and severity assessment of MRPs are discussed in chapter 8, and the univariable relationship between the candidate predictors and outcome events (patients with at least one MSP MRP) is discussed in chapter 9.

6.4.1 Flow of patients through the study

One thousand six hundred and fifty two patient admissions were included in the study, 1,100 from Hospital A and 552 from Hospital B, as summarised in Figure 3**.** Of these admissions, 149 (9%) were excluded (114 did not meet the eligibility criteria, and prescribing records and/or medical records were unavailable for 35).

Of the remaining 1,503 study admissions, 1,378 were followed up until discharge from hospital; 933 (93%) of 1,006 at Hospital A, and 445 (90%) of 497 at Hospital B. The remaining admissions were followed until the study end date (two weeks after inclusion of the final patient at each site). Eight hundred and ninety four (59.5%) experienced at least one MRP, with 610 (40.6%) experiencing the outcome event, namely at least one MSP MRP.

Fifty seven patients entered the study twice, 46 at Hospital A and 11 at Hospital B. One further patient at Hospital B entered the study three times. The total number of patients included in the study was therefore 1,444 (960 at Hospital A and 484 at Hospital B).



MRP = medication related problem

Figure 3 – Participant flow diagram

6.4.2 Characteristics of participants

The key characteristics of the 1,503 admissions included in the regression analyses are summarised in Table 17. This includes demographic variables and all candidate predictors.

The median age of participants was 75 years (range 17 to 103), with female patients comprising 693 (46.1%) of 1,503 admissions. The statistical analysis suggests significant differences between study sites in terms of socioeconomic status, with Hospital A having a lower median socioeconomic status ($p < 0.001$). Hospital A also had a lower proportion of patients with an ethnic origin described as 'white', suggesting greater ethnic diversity ($p < 0.001$). A difference was also observed in the proportion of admissions with allergies ($p = 0.002$), and in antimicrobial use ($p = 0.0073$), although these differences were not statistically significant given the Bonferroni corrected p value of 0.0012.

The laboratory results at the two sites were broadly comparable, although the median platelet count was lower at Hospital B compared to Hospital A ($p < 0.001$). This may be explained by the of lower standard reference range at Hospital B (Hospital A: 150-450 10^9/L, Hospital B: 120-400 10^9/L). In addition one needs to consider the clinical significance of the difference, as the median results for both sites are within the standard reference ranges; this means that while the difference may be statistically significant, it may not be considered to be of clinical significance. This issue was discussed with Dr Li Wei (academic supervisor), and it was agreed that it would be appropriate to combine the data for the regression analyses irrespective of the difference, as this has the potential to increase the generalisability of the MOAT (as reference ranges often differ between hospitals). This also avoided the need to categorise the data, which as discussed in chapter 3, is associated with reduced model reliability[60].

Table 17 – Characteristics of study admissions

Characteristic	Hospital A (admissions = 1,006) n (% of admissions)	Hospital B (admissions = 497) n (% of admissions)	All patients (admissions = 1,503) n (% of admissions)	*p* value (test for difference between study sites)
Demographic				
Age (years)	Median: 75 IQR: 58-85 Range: 19-103	Median: 76 IQR: 57.5-86 Range: 17-100	Median: 75 IQR: 58-85	0.358 (Mann-Whitney)
Gender (female)	446 (44.3)	247 (49.7)	693 (46.1)	0.050 (Chi-square)
Socioeconomic status*, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]†	Median: 40 IQR: 22-67 Range: 3-100	Median: 72 IQR: 49-91 Range: 10-100	Median: 50 IQR: 30-79	<0.001 (Mann-Whitney)
Ethnic origin (white)*	777 (82.3)	431 (93.1)	1208 (85.9)	<0.001 (Chi-square)
Patient related				
Previous allergy*	362 (36.0)	220 (44.3)	582 (38.8)	0.0020 (Chi-square)
Body mass index* (kg/m^2)	Median: 24.7 IQR: 21.4-29.5 Range: 10.6-65.5	Median: 25.2 IQR: 21.5-28.7 Range: 14.9-55.6	Median: 24.9 IQR: 21.4-29.1	0.852 (Mann-Whitney)
Number of hospital admissions in previous 6 months	Median: 0 IQR: 0-1 Range: 0-10	Median: 0 IQR: 0-1 Range: 0-7	Median: 0 IQR: 0-1	0.956 (Mann-Whitney)
Primary diagnosis:				
Endocrine and metabolic	52 (5.2)	30 (6.0)	82 (5.5)	0.486 (Chi-square)
Nervous system and mental disorders	92 (9.1)	57 (11.5)	149 (9.9)	0.156 (Chi-square)
Cardiovascular system	225 (22.4)	90 (18.1)	315 (21.0)	0.056 (Chi-square)
Respiratory system	219 (21.8)	113 (22.7)	332 (22.1)	0.671 (Chi-square)
Gastrointestinal system	89 (8.9)	55 (11.1)	144 (9.6)	0.169 (Chi-square)
Genitourinary system	100 (9.9)	44 (8.9)	144 (9.6)	0.501 (Chi-square)
Musculoskeletal-intergumentary systems	68 (6.8)	25 (5.0)	93 (6.2)	0.191 (Chi-square)
All other categories	161 (16.0)	83 (16.7)	244 (16.2)	0.731 (Chi-square)
Number of comorbidities	Median: 3 IQR: 2-5 Range: 0-13	Median: 4 IQR: 2-5 Range: 0-11	Median: 4 IQR: 2-5	0.845 (Mann-Whitney)
History of dementia	103 (10.2)	58 (11.7)	161 (10.7)	0.399 (Chi-square)
Length of hospital stay (days)	Median: 5 IQR: 2-13 Range: 0-148	Median: 5 IQR: 2-11 Range: 0-179	Median: 5 IQR: 2-12	0.110 (Mann-Whitney)
Medicines related				
Medicines reconciliation completed	810 (80.5)	482 (97.0)	1292 (86.0)	<0.001 (Chi-square)
Number of medicines	Median: 7 IQR: 5-10 Range: 0-27	Median: 8 IQR: 5-11 Range: 0-21	Median: 8 IQR: 5-10	0.095 (Mann-Whitney)
Parenteral administration	674 (67.0)	334 (67.2)	1008 (67.1)	0.937 (Chi-square)

Continued from previous page…

Characteristic	Hospital A (admissions = 1,006) n (% of admissions)	Hospital B (admissions = 497) n (% of admissions)	All patients (admissions = 1,503) n (% of admissions)	*p* value (test for difference between study sites)
Use of high-risk medicines:				
Anticoagulants	194 (19.3)	118 (23.7)	312 (20.8)	0.045 (Chi-square)
Therapeutic heparin	152 (15.1)	70 (14.1)	222 (14.8)	0.598 (Chi-square)
Anti-diabetic medication	210 (20.9)	89 (17.9)	299 (19.9)	0.175 (Chi-square)
Opiates	100 (9.9)	45 (9.1)	145 (9.6)	0.584 (Chi-square)
Aminoglycosides and glycopeptides	58 (5.8)	47 (9.5)	105 (7.0)	0.083 (Chi-square)
Other antimicrobials	657 (65.3)	280 (56.3)	937 (62.3)	0.0073 (Chi-square)
Theophylline and aminophylline	28 (2.8)	10 (2.0)	38 (2.5)	0.370 (Chi-square)
Epilepsy medicines	161 (16.0)	66 (13.3)	227 (15.1)	0.165 (Chi-square)
Antipsychotics (excluding clozapine)	56 (5.6)	36 (7.2)	92 (6.1)	0.202 (Chi-square)
Immunosuppressants	12 (1.2)	9 (1.8)	21 (1.4)	0.337 (Chi-square)
Cytotoxics	11 (1.1)	3 (0.6)	14 (0.9)	0.355 (Fisher's exact)
Lithium	4 (0.4)	2 (0.4)	6 (0.4)	1.000 (Fisher's exact)
Antiarrhythmics	110 (10.9)	40 (8.0)	150 (10.0)	0.079 (Chi-square)
Antidepressants	239 (23.8)	112 (22.5)	351 (23.4)	0.598 (Chi-square)
Other (clozapine, anti-retrovirals, medicines for Parkinson's disease)	27 (2.7)	13 (2.6)	40 (2.7)	0.938 (Chi-square)
Laboratory results				
Renal function - estimated glomerular filtration rate*‡ (ml/min/1.73m^2)	Median: 73 IQR: 53-95 Range: 5-294	Median: 73 IQR: 51-97 Range: 3-309	Median: 73 IQR: 53-99	0.871 (Mann-Whitney)
Liver disease	107 (10.6)	57 (11.5)	164 (10.9)	0.626 (Chi-square)
Serum albumin* (g/L)	Mean: 33.0 SD: 5.8 Range: 10-51	Mean: 32.9 SD: 6.4 Range: 7-55	Mean: 33.0 SD: 6.0	0.791 (Two-sample t-test)
Serum potassium* (mmol/L)	Mean: 4.5 SD: 0.63 Range: 2.3-7.8	Mean: 4.4 SD: 0.59 Range: 2.7-6.9	Mean: 4.4 SD: 0.62	0.017 (Two-sample t-test)
Serum sodium* (mmol/L)	Mean: 137.1 SD: 5.1 Range: 111-170	Mean: 137.4 SD: 5.5 Range: 113-165	Mean: 137.2 SD: 5.2	0.238 (Two-sample t-test)
White cell count* (10^9/L)	Median: 9.8 IQR: 7.5-13.0 Range: 0.3-93.0	Median: 9.8 IQR: 7.7-12.6 Range: 2.5-32.9	Median: 9.8 IQR: 7.5-12.8	0.707 (Mann-Whitney)
Median platelet count* (10^9/L)	Median: 249 IQR: 196-320 Range: 5-977	Median: 230 IQR: 185-293 Range: 14-738	Median: 244 IQR: 192-312	<0.001 (Mann-Whitney)

* For patients without missing data (further details provided in Appendix A6.6)

† Deprivation rank based on patients' postcode, shown as the ranked position as a percentage of all neighbourhoods in England (where 1 is the most deprived)

‡ Glomerular filtration rate estimated using modified Modification of Diet in Renal Disease equation[121]

IQR = interquartile range, SD = standard deviation
Bonferroni adjusted *p* value used to judge statistical significance 0.0012 (based on 43 statistical tests)

6.4.3 Analysis of missing data

Thirty five (2.1%) of the 1,652 patient admissions were excluded from the study as prescribing and/or medical records were unavailable (Figure 3). These were excluded to ensure complete data on medicine-related candidate predictors were available for the regression analyses. Thirty four of these admissions occurred at Hospital B, where paper-based medical and prescribing records were used. Of these, the medical records could not be located for seven admissions, and the study admission and/or prescribing records were missing for the remaining 27. This issue was raised with the Medical Records department, who advised that the records were likely to be misfiled. Prescribing records were unavailable for one patient at Hospital A. This appeared to be due to the patient's unique identification number being changed during the admission.

Of the 1,503 admissions included in the regression analyses, 449 (29.9%) had one or more missing data point. Table 18 gives the number of admissions with missing values, and number of values missing. There was no evidence for a difference in the proportion of admissions with missing data between study sites (p = 0.830).

Table 18 – Number of missing values per admission

Number of missing values	**Number of admissions (admissions = 1,503)** n (% of all admissions)
0	1054 (70.1)
1	390 (26.0)
2	52 (3.5)
3	4 (0.3)
6	1 (0.07)
7	2 (0.1)

Appendix A6.6 gives a breakdown of the number of missing data points for each study variable. This shows that no data were missing for age, gender, primary diagnoses, comorbidities, number of previous hospital admissions, length of hospital stay, liver disease, and medicine usage.

Socioeconomic status was unavailable for six admissions. This was because the English Indices of Deprivation 2015[149] is based on patients' postcodes, and the six patients with missing socioeconomic data were either not resident in England (four

admissions), or had no fixed abode (two admissions). Allergy status was not available for one admission, although this patient was prescribed nine medicines during their admission, including antimicrobials, which may suggest they had no allergies (as nursing staff at the study sites are required to check a patient's allergy status before administering any medication). Body mass index (BMI) was the variable with the highest number of missing values, with data missing for 341 (22.7% of admissions). Of these, 98 admissions had a weight measurement only, 96 had height only, and 147 had neither. Although it is not possible to ascertain the reason(s) why these data were missing, there are numerous possibilities including patients being acutely unwell (therefore unable to stand / sit for measurement to take place), short length of hospital stay (i.e. patient discharged before measurement could be taken), or a refusal to be weighed / measured. It was not possible to estimate the renal function for nine admissions (0.6%) as serum creatinine measurements were unavailable. In six cases this appears to be due to high bilirubin levels interfering with the creatinine assay[150]. Similarly, of the 30 missing serum potassium values, 27 were missing due to the laboratory being unable to analyse the sample (rather than the test not being requested), which may be due to haemolysis of the blood sample[151]. Of the remaining 43 missing laboratory results (albumin, sodium, white cell count and platelet count), 38 were not tested, and five were tested but the sample could not be analysed (two for white cell count and three for platelet count).

Ethnic origin data were collected for descriptive purposes rather than for use in the regression modelling, but were unavailable for 96 admissions. Reasons for the missing data could include patients being too unwell to provide the information, having a complex racial heritage (although the study sites used categories for mixed heritage and 'other' ethnicity), or a reluctance to disclose this information. It was necessary to establish ethnic origin to estimate renal function, but only to the extent of whether patients were 'black' or 'non-black' (discussed in section 6.2.4.1).

To inform the possible missingness mechanism I compared the characteristics of admissions with missing values and those with completely observed data (Appendix A6.7). Statistically significant differences were observed, with results suggesting that fewer patients with missing values experienced the outcome event ($p < 0.0001$). Patients with missing values were also statistically: younger ($p < 0.001$), had fewer admissions in the previous six months ($p < 0.001$), fewer comorbidities ($p < 0.001$), fewer medicines ($p < 0.001$), and a shorter length of hospital stay ($p < 0.001$). In addition, a lower proportion of patients with missing data received parenteral medicines

($p < 0.001$), and a higher proportion were male ($p = 0.0015$). There was also weak evidence (given the Bonferroni adjusted p value of 0.0028) that a higher proportion of patients with missing values had results within the standard reference range for serum albumin ($p = 0.0069$), normal renal function ($p = 0.0334$), and a higher BMI ($p = 0.029$). These findings may suggest that patients with missing data had fewer indicators of long-standing illness burden, and/or severity of the current admission compared to those with completely observed data. It is of note that some markers of current illness (deranged serum potassium, sodium, white cell count, and platelets) were comparable between the two groups. The slightly higher BMI in the group with missing data may suggest that patients with higher BMIs were less likely to have their weight and/or height recorded.

Missingness mechanisms fall into three categories, missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR)[148], which have been described as follows[147]:

- data are MCAR if the probability that a particular observation is missing does not depend on the value of any observed variable(s);
- data are MAR if, given the observed data, the probability that observations are missing is independent of the actual value of the missing data;
- data are MNAR if the probability of missing still depends on the missing value even after taking the available data into account.

It is known that differences in the extent and type of missing data, and the methods used to handle this missing data, may greatly influence model development and predictive performance of prognostic models[60]. It was therefore necessary to consider the possible missingness mechanism to decide how to deal with the missing data. Where data are MCAR, it is considered reasonable to use complete-case analysis, as participants with missing data are likely to be a truly random subset of the study sample[60]. Where missing data are related to other observed participant data (i.e. MAR), a complete-case analysis would lead to a non-random subset, and biased results[60 147]. Multiple imputation is generally considered to be the preferred method to handle data that are MAR[60 147], but is not appropriate where data are MNAR, as MNAR data are related to unobserved data, and therefore cannot be plausibly estimated from the observed study variables.

Of the MOAT study data, the patients excluded due to the unavailability of prescribing and/or medical records were likely to be MCAR, as the misfiling was likely to be

completely random, and unrelated to any patient related factors. Exclusion of these patients is therefore unlikely to bias the study results.

Of the variables included in MOAT development, the comparison of admissions with and without missing data suggests the missingness mechanism was MAR rather than MCAR, as admissions with missing data were clearly not a random subset of the study sample. On this basis it is appropriate to use multiple imputation, but as it is not possible to distinguish between MAR and MNAR using observed data, sensitivity analyses will be required to investigate possible departure from the MAR assumption[94] [147]. The imputation of missing data and sensitivity analyses will be discussed further in chapter 9.

6.4.4 Data entry checks for candidate predictors

Data entry checks were performed for 152 admissions (99 from Hospital A and 53 from Hospital B), giving a total of 2,432 data item checks. Accuracy was calculated as the percentage of data items recorded correctly. Results are summarised in Table 19.

Table 19 – Accuracy of data entry for candidate predictors

Hospital	Prescribing records		Laboratory reports		Patient medical records	
	Total items checked	Accuracy n (% of items checked)	Total items checked	Accuracy n (% of items checked)	Total items checked	Accuracy n (% of items checked)
Hospital A (99 admissions)	396	396 (100)	891	891 (100)	297	297 (100)
Hospital B (53 admissions)	212	212 (100)	477	474 (99.4)*	159	159 (100)

* Accuracy prior to remedial action

As shown in Table 19, the accuracy of data entry for the randomly selected admissions at Hospital A was 100%, as were the prescribing, and patient medical record checks at Hospital B.

Data entry checks of laboratory reports for the first eight randomly selected admissions at Hospital B (a total of 72 data items) identified one numerical transcription error. I therefore checked all laboratory data items collected to that point (91 admissions, 819 data items), and found one further error. This was discussed with the data analyst and the data extraction method refined. No further numerical transcription errors were identified in the remaining 45 randomly selected admissions (405 data items), but two further errors were identified where results were recorded as ‘not tested’, despite being

available. I therefore checked all results recorded as 'not tested'. This gave an overall error rate for the randomly selected sample of three (0.6%) of 477 data items (95% confidence interval -0.1 to 1.3%), which I considered to be acceptable, particularly given the remedial action and further checks undertaken.

6.4.5 MRP descriptive data

The MRP descriptive data results are presented in three sections: MRP identification, categorisation (using Basger's classification system[114]), and the prevalence of MRPs in the study sample.

This chapter includes data for all MRPs irrespective of severity or preventability. Results for MSP MRPs are presented in chapter 8.

6.4.5.1 MRP identification

A total of 2,747 MRPs were reported for the 1,503 study admissions. Eight MRPs (all from Hospital A) were identified via the hospital incident reporting system; these occurred and were resolved outside of pharmacy working hours. The remainder were reported by pharmacy staff after being identified personally or through discussion with other healthcare professionals.

As shown in Table 20, 11 MRPs were excluded from the analysis as they were duplicate entries (i.e. the same MRP reported twice on the same day). The remaining 2,736 were reviewed by the expert panel, and 122 (4.4%) were not considered to be true MRPs. There was no evidence for a statistically significant difference in the proportion of MRPs excluded due to non-validation by expert panel members between the study sites (p = 0.129).

Table 20 – Medication related problem identification

	Medication related problems (MRPs)		
	Hospital A n (% of MRPs)	**Hospital B** n (% of MRPs)	**All patients** n (% of MRPs)
Number of MRPs reported	1,778 (100)	969 (100)	2,747 (100)
Number of MRPs excluded as duplicate reports	10 (0.6)	1 (0.1)	11 (0.4)
Number of MRPs excluded as not considered to be true MRPs by expert panel	71 (4.0)	51 (7.3)	122 (4.4)
Total MRPs remaining	**1,697 (95.4)**	**917 (94.6)**	**2,614 (95.2)**

Table 21 gives the reasons for non-validation of the 122 excluded MRPs (i.e. why they were not considered to be true MRPs). This shows that 'unnecessary pharmacy contribution' formed the largest category, accounting for 50 (41%).

Table 21 – Reasons for non-validation of medication related problems

Reason for non-validation	Non-validated medication related problems (MRPs)		
	Hospital A (non-validated MRPs = 71) n (% of non-validated MRPs)	**Hospital B** (non-validated MRPs = 51) n (% of non-validated MRPs)	**All patients** (non-validated MRPs = 122) n (% of non-validated MRPs)
Lack of documentation (rather than true MRP) e.g. indication / duration for medicine not stated (rather than inappropriate use)	9 (12.7)	3 (5.9)	12 (9.8)
Unnecessary pharmacy contribution e.g. pharmacist advised use of once daily (modified release) oral nitrate rather than twice daily 'to get better serum concentrations' (when both were considered by panel members to be clinically acceptable)	36 (50.7)	14 (27.5)	50 (41.0)
Local policy issue only e.g. formulary issues	19 (26.8)	4 (7.8)	23 (18.9)
Non-clinically significant drug interaction e.g. concomitant use of omeprazole and clopidogrel (although there is a theoretical drug interaction, panel members felt it was not clinically significant)	7 (9.9)	22 (43.1)	29 (23.8)
Insufficient information to permit validation (but from data available appears to be 'an unnecessary pharmacy contribution' and/or minor severity)* e.g. anxiolytic not prescribed on discharge, which would have been appropriate if treatment intended during hospitalisation only (but no further information given)	0	8 (15.7)	8 (6.6)

* Due to a delay between MRP data collection and expert panel consensus discussions (approximately one year due to temporary unavailability of one panel member) it was not possible to seek further information/clarification from reporting pharmacists

A review of the eight MRPs with 'insufficient information to permit validation' found they related to seven study admissions. Of these admissions, two patients experienced another MRP that met the requirements for a MSP MRP. The remaining five patients were recorded as not having experienced an MRP/MSP MRP.

Additional MRP descriptive data are summarised in Table 22. This shows that the majority of MRPs (93.2%) were identified during routine ward visits, and resolved by pharmacy staff (97.9%). MRPs were also more frequently identified during (or before)

the first review of the patient (74.6%); this may be related to the high proportion of medicines reconciliation related MRPs (55.9%), as medicines reconciliation is often undertaking during the first ward review[5].

Table 22 – Descriptive data for medication related problems

	Hospital A (MRPs = 1,697) n (% of MRPs)	**Hospital B (MRPs = 917)** n (% of MRPs)	**All patients (MRPs = 2,614)** n (% of MRPs)	***p* value (test for difference between study sites)**
When MRP identified:				
During ward visit	1,584 (93.3)	853 (93.0)	2,437 (93.2)	0.455 (Chi-square)
In the pharmacy department	105 (6.2)	56 (6.1)	161 (6.2)	
Other (pharmacist referral / reported via hospital incident reporting system)	8 (0.5)	8 (0.9)	16 (0.6)	
Who resolved MRP:				
Pharmacy staff	1,650 (97.2)	908 (99.0)	2,558 (97.9)	0.0026 (Chi-square)
Other healthcare professionals	47 (2.8)	9 (1.0)	56 (2.1)	
Stage in patient stay MRP identified:				
During first ward review (or before)	1246 (73.4)	704 (76.8)	1,950 (74.6)	0.0036 (Chi-square)
Remainder of inpatient stay	307 (18.1)	119 (13.0)	426 (16.3)	
Clinical screening at discharge	144 (8.5)	88 (9.6)	232 (8.9)	
Missing data	0	6 (0.7)	6 (0.2)	
Medicines reconciliation discrepancy	935 (55.1)	526 (57.4)	1,461 (55.9)	0.260 (Chi-square)

MRP = medication related problem

Bonferroni adjusted *p* value used to judge statistical significance 0.0125 (based on 4 statistical tests)

There was evidence for a statistically significant difference between study sites for the stage in patient stay when MRP identified. The results suggest that a greater proportion of MRPs were identified during the first ward review at Hospital B compared to Hospital A (76.8% compared to 73.4%), with fewer identified during the remainder of the admission (13.0% compared to 18.1%). This may be explained by the higher number of patients receiving medicines reconciliation at Hospital B (Table 17, 97% compared to 80.5%, p <0.0001), as this may have led to a greater proportion of pharmacists' time being spent reviewing new rather than existing patients. It is not possible to investigate this hypothesis further as no data were collected on the frequency or duration of pharmacy reviews.

There is also evidence that the proportion of MRPs resolved by pharmacy staff (as opposed to other healthcare professionals) was lower at Hospital A compared to Hospital B (97.2% and 99.0% respectively, p = 0.0026). While this is of limited practical significance (given the relatively small percentage difference), it may be

explained in part by the eight MRPs identified via the hospital incident reporting system at Hospital A, as these were resolved without pharmacist intervention, and the number of MRPs related to non-availability of medicines that were resolved by nursing staff (21 at Hospital A, compared to three at Hospital B).

MRPs identified when clinically screening discharge prescriptions were subdivided (to provide additional information on workflow, and to potentially inform implementation of the MOAT into clinical practice) as follows:

- those occurring prior to the discharge prescription being written (i.e. MRPs that had already occurred but not identified / resolved, such as medicines reconciliation omissions);
- those occurring when the discharge prescription was written (e.g. transcription errors, technical errors, MRPs related to newly prescribed medication, or alterations made at the point of hospital discharge).

The results are summarised in Table 23, and show that the majority of discharge related MRPs occurred when the discharge prescription was written (70.4%). There was no statistically significant difference between study sites (p = 0.126).

Table 23 – Breakdown of identification of medication related problems identified when discharge prescription screened

When medication related problem (MRP) occurred	**Hospital A (MRPs at discharge = 145)** n (% of MRPs at discharge)	**Hospital B (MRPs at discharge = 88)** n (% of MRPs at discharge)	**All patients (MRPs at discharge = 233)** n (% of MRPs at discharge)	***p* value (test for difference between study sites)**
Prior to writing discharge prescription	48 (33.1)	21 (23.9)	69 (29.6)	0.126 (Chi-square)
When discharge prescription written	97 (66.9)	67 (76.1)	164 (70.4)	

6.4.5.2 MRP classification

The classification of the 2,614 MRPs is summarised in Table 24. This shows that the most frequently identified subcategory was ‘indication not treated / missing therapy’, accounting for 1,116 (42.7%). MRPs related to dose selection were the next most frequently reported with ‘dose too low’ and ‘dose too high’ each accounting for 9.6% of the total MRPs (252 and 250 respectively).

Table 24 – Classification of medication related problems

Medication related problem (MRP) subcategory	**Hospital A (MRPs = 1,697)** n (% of MRPs)	**Hospital B (MRPs = 917)** n (% of MRPs)	**All patients (MRPs = 2,614)** n (% of MRPs)	***p* value (test for difference between study sites)**
1. Drug selection				
1.1 Inappropriate drug	85 (5.0)	36 (3.9)	121 (4.6)	0.210 (Chi-square)
1.2 No indication for drug / duplication	73 (4.3)	42 (4.6)	115 (4.4)	0.740 (Chi-square)
1.3 Interaction (drug-drug, or drugs and food / alcohol)	19 (1.1)	16 (1.8)	35 (1.3)	0.184 (Chi-square)
1.4 Indication not treated / missing therapy	677 (39.9)	439 (47.9)	1,116 (42.7)	0.000083 (Chi-square)
1.5 More cost effective drug available	8 (0.5)	0	8 (0.3)	0.057 (Fisher's exact)
1.6 Synergistic / preventive drug required and not given	25 (1.5)	2 (0.2)	27 (1.0)	0.0025 (Chi-square)
2. Drug form				
2.1 Inappropriate or suboptimal drug form	59 (3.5)	22 (2.4)	81 (3.1)	0.130 (Chi-square)
3. Dose selection				
3.1 Drug dose too low	166 (9.8)	86 (9.4)	252 (9.6)	0.739 (Chi-square)
3.2 Drug dose too high	159 (9.4)	91 (9.9)	250 (9.6)	0.646 (Chi-square)
3.3 Dosage regimen not frequent enough	5 (0.3)	4 (0.4)	9 (0.3)	0.728 (Fisher's exact)
3.4 Dosage regimen too frequent	9 (0.5)	2 (0.2)	11 (0.4)	0.348 (Fisher's exact)
3.5 Dose needs adjustment to organ function or change in disease state	22 (1.3)	8 (0.9)	30 (1.2)	0.331 (Chi-square)
3.6 Dosage instructions unclear, incomplete or not understood by patient / carer*	N/A*	N/A*	N/A*	N/A*
4. Treatment duration / withdrawal				
4.1 Duration of treatment too short	4 (0.2)	1 (0.1)	5 (0.2)	0.663 (Fisher's exact)
4.2 Duration of treatment too long	39 (2.3)	15 (1.6)	54 (2.1)	0.256 (Chi-square)
4.3 Inappropriate abrupt withdrawal†	2 (0.1)	0	2 (0.08)	0.544 (Fisher's exact)
5. Drug use process				
5.1 Inappropriate timing of administration / dosing by prescriber; administration error by nurse	58 (3.4)	20 (2.2)	78 (3.0)	0.076 (Chi-square)
5.2 Drug underused / under-administered	11 (0.7)	11 (1.2)	22 (0.8)	0.141 (Chi-square)
5.3 Drug overused / over-administered	0	1 (0.1)	1 (0.04)	0.351 (Fisher's exact)
5.4 Drug not taken / administered at all	7 (0.4)	11 (1.2)	18 (0.7)	0.020 (Chi-square)
5.5 Wrong drug taken by patient	0	0	0	N/A
5.6 Drug abused	0	0	0	N/A
5.7 Patient or nurse uses drug incorrectly through lack of knowledge or barriers (e.g. swallowing, dexterity)	1 (0.1)	2 (0.2)	3 (0.1)	0.283 (Fisher's exact)

Continued from previous page…

Medication related problem (MRP) subcategory	**Hospital A (MRPs = 1,697)** n (% of MRPs)	**Hospital B (MRPs = 917)** n (% of MRPs)	**All patients (MRPs = 2,614)** n (% of MRPs)	***p* value (test for difference between study sites)**
5.8 Adequate information not provided or not understood or misunderstood or not followed*	N/A*	N/A*	N/A*	N/A*
5.9 Drugs stored inappropriately / expired drug administered / preparation error	0	3 (0.3)	3 (0.1)	0.043 (Fisher's exact)
6. Logistics				
6.1 Prescribed drug not available	38 (2.2)	13 (1.4)	51 (2.0)	0.148 (Chi-square)
6.2 Drug order incorrect, incomplete, poorly legible / illegible / illegal / incorrect / allergy status incomplete	126 (7.4)	75 (8.2)	201 (7.7)	0.490 (Chi-square)
6.3 Error in drug selection	91(5.4)	15 (1.6)	106 (4.1)	0.00004 (Chi-square)
7. Monitoring				
7.1 Monitoring too frequent	1 (0.1)	0	1 (0.04)	1.000 (Fisher's exact)
7.2 No or too infrequent monitoring	6 (0.4)	1 (0.1)	7 (0.3)	0.433 (Fisher's exact)
7.3 Inappropriate test ordered	1 (0.1)	0	1 (0.04)	1.000 (Fisher's exact)
7.4 Patient unable to attend / pay for monitoring*	N/A*	N/A*	N/A*	N/A*
8. Unexpected reaction / adverse drug reaction (ADR) / no obvious cause				
8.1 An ADR occurred	5 (0.3)	1 (0.1)	6 (0.2)	0.672 (Fisher's exact)
8.2 No obvious cause of treatment failure	0	0	0	N/A

* Category not used for MOAT study as relates to primary care (discussed in section 6.2.3)

† Category not included in Basger's original classification system[114] (discussed in section 6.2.3)

N/A = not applicable

Bonferroni adjusted *p* value used to judge statistical significance 0.0018 (based on 28 statistical tests)

Statistically significant differences between study sites were found for the following MRP subcategories (using Bonferroni corrected *p* value of 0.0018):

- 'indication not treated / missing therapy' ($p = 0.000083$);
- 'error in drug selection' ($p = 0.00004$).

Of the MRPs categorised as 'indication not treated / missing therapy' 91.2% were medicines reconciliation related discrepancies, 624 (92.2%) of the 677 MRPs at Hospital A, and 394 (89.8%) of 439 MRPs at Hospital B. Although there was no evidence for a difference in the proportion of medicines reconciliation discrepancies

between sites (shown in Table 22), there does appear to be a difference in the type of errors identified during medicines reconciliation, with 'indication not treated / missing therapy' accounting for significantly more of the medicines reconciliation related discrepancies at Hospital B compared to Hospital A, that is, 394 (74.9%) of the 526 medicines reconciliation related discrepancies at Hospital B, and 624 (66.7%) of the 935 discrepancies at Hospital A (p = 0.0012).

The difference between study sites in the proportion of MRPs classified as 'error in drug selection' may be explained by the different medication prescribing systems used, with Hospital A using electronic prescribing, and Hospital B using a paper-based system. Previous research suggests that electronic prescribing systems can reduce the overall occurrence of prescribing errors but can also contribute to new types of errors, including selection errors related to the use of 'drop-down' lists[152 153]. Possible examples of selection errors that may be related to the electronic prescribing system at Hospital A include ten MRPs where soluble, effervescent, or orodispersbile formulations of medicines were prescribed where a standard formulation appeared to be intended, and 16 where an incorrect inhaler device was selected. Some selection errors were potentially more clinically significant, including a patient prescribed Levonelle (an emergency hormonal contraceptive) when levetiracetam (a medicine for epilepsy) was intended.

There was weak evidence for a difference between sites for the category 'synergistic / preventive drug required and not given' (p = 0.0025), which may be explained by the routine use of probiotic supplements at Hospital A only (for older patients prescribed broad spectrum antibiotics). This accounted for 16 of the 25 MRPs at Hospital A, which if excluded from the analysis, results in no evidence for a statistically significant difference (Fisher's exact test, p = 0.348).

6.4.5.3 Prevalence of MRPs

As shown in Table 25, 894 (59.5%) of 1,503 study admissions experienced at least one MRP. Of the admissions that experienced an MRP, the number of MRPs per admission ranged from one to 17, with a median of two.

Table 25 – Prevalence of medication related problems

Medication related problem (MRP) characteristics	**Hospital A (admissions = 1,006)** n (% of admissions)	**Hospital B (admissions = 497)** n (% of admissions)	**All patients (admissions = 1,503)** n (% of admissions)	***p* value (test for difference between study sites)**
Admissions with at least one MRP*	576 (57.3)	318 (64.0)	894 (59.5)	0.012 (Chi-square)
Number of MRPs per admission (of the admissions with an MRP)	Median: 2 IQR: 1-4 Range: 1-16	Median: 2 IQR: 1-4 Range: 1-17	Median: 2 IQR: 1-4	0.390 (Mann-Whitney)

* MRPs considered by expert panel to be true MRPs

Bonferroni adjusted *p* value used to judge statistical significance 0.025 (based on 2 statistical tests)

There was evidence for a difference in the proportion of admissions who experienced at least one MRP between study sites ($p = 0.012$). This may be explained by the higher proportion of patients receiving medicines reconciliation at Hospital B (97% versus 80.5% at Hospital A, as shown in Table 17) given the high number of unintentional discrepancies known to occur at transitions of care[37].

6.5 Discussion

Key findings

Of the 1,652 patient admissions included in the study, 1,503 met the study's eligibility criteria and were included in the analyses. Over 2,700 MRPs were reported for the eligible admissions, of which 2,614 were considered to be true MRPs by an expert panel. This equates to an average of 1.7 MRPs per admission. The most frequently identified MRP subcategories were 'indication not treated / missing therapy', 'dose too low' and 'dose too high', accounting for almost 62% of MRPs. Eight hundred and ninety four (59.5%) of the eligible admissions experienced at least one MRP. As anticipated, there were significant differences in socioeconomic status and ethnicity between the two study sites.

Comparison with previous literature

The MRP prevalence found in the present study is consistent with findings from other research. Blix *et al*[78] (Norway, 2004) reported that 81% of hospitalised patients (from internal medicine and rheumatology wards) experienced an MRP, with an average of 2.1 clinically relevant MRPs recorded per patient. In comparison, the present study found that 59.5% of admissions experienced at least one MRP, with an average of 1.7 MRPs per admission. The data collection method and sample population were similar for both studies, but the prevalence reported by Blix *et al* may have been higher due to:

- higher prevalence of ADRs, occurring in 7.8% of the 827 patients (number not reported), whereas only 6 ADRs were reported for the 1,503 MOAT patients;
- higher proportion of patients reported by Blix *et al* to have 'medical chart errors' (16.3%). This category accounted for only 7.7% of MOAT MRPs;
- inclusion of the MRP subcategory 'information / therapy discussions', which was reported for 17.3% of patients by Blix *et al*. This category does not form part of Basger's classification system[114] (as used for the MOAT study).

Two, more recent studies (both using similar data collection methods and target populations to the MOAT study), reported MRP prevalence closer to that found by the MOAT study. Wilmer *et al* (Netherlands, 2015)[90], reported that 70 (53%) of 131 patients experienced one or more MRP, with an average of 1.0 MRPs per patient. Ayalew *et al* (Ethiopia, 2015)[72], reported MRPs in 117 (52%) of 225 study subjects, and an average of 0.68 MRPs per patient.

Urbina *et al* (Spain, 2014)[52] conducted a larger study, comprising 8,713 admissions; of these, 2,425 (27.8%) experiencing at least one MRP, with an average of 0.3 MRPs per

admission. This prevalence is significantly lower than the studies discussed above, but Urbina's results are not directly comparable as surgical and maternity patients were included (in additional to general medicine), and a computerised warning system used to identify MRPs (rather than identification by pharmacy staff).

Due to the use of different MRP classification systems it is not possible to directly compare the distribution of MRP subcategories between these studies, but it appears that the use of different data collection methods may have led to some variation. For example, Ayalew *et al* used a 'drug interaction checker' as part of their MRP identification process, which may have resulted in the high proportion of drug interactions (48%). Similarly, Urbina *et al* used a computerised system to identify MRPs, and found that 'incorrect use of the computerised physician order entry system' accounted for 23.9% of MRPs.

Despite the potential differences in MRP sub-categorisation, there do appear to be some similarities between the findings of the MOAT study and previous research. The MOAT study identified 'indication not treated / missing therapy' as the highest MRP subcategory, accounting for 39.9%, which is comparable with Wilmer *et al*, where 'under-treatment' accounted for 35.5%. Wilmer *et al* also reported that incorrect dosing (overdose or under-dose) accounted for 25.0%, which is similar to the 19.1% found in the MOAT study.

Ayalew *et al* and Urbina *et al* both reported a higher proportion of 'drug interactions' than the MOAT study (46.0%, 11.2%, and 1.3% respectively), which may be due to the methods of identification used (a drug interaction checker by Ayalew *et al*, computerised system by Urbina *et al*, and pharmacists' clinical judgement for the MOAT study). Although not specified, it is possible that Ayalew and Urbina *et al* included all potential drug interactions irrespective of clinical relevance, as opposed to the MOAT study, where only drug interactions considered to be clinically significant were counted as true MRPs.

Strengths and limitations

A strength of the approach taken for MOAT data collection was adherence with PROGRESS[53 55], TRIPOD[94], and CHARMS[60] recommendations. This has the potential to enhance the quality of data collection, reduce bias, and facilitate full and detailed reporting of results to permit the quality and relevance of the study to be adequately assessed. Other strengths include the detailed description of data collection methods (to ensure reproducibility of results), and the choice of candidate predictor definitions

that are clinically relevant and straight forward to use, so enhancing the potential clinical credibility and usability of the MOAT in clinical practice.

A limitation is the possible underestimation of the prevalence of MRPs due to:

- not all participants receiving medicines reconciliation - this may have resulted in non-identification of MRPs, therefore measurement bias[154];
- potential differences in MRP identification depending on the knowledge, experience and skills of pharmacists collecting data. A simulated 'MRP identification assessment exercise' was therefore used to permit this to be quantified (described in chapter 7);
- the observational nature of the study, which meant data collection was not carried out under strict trial conditions, with pharmacy staff required to complete data collection in addition to other routine duties, and ward staff required to inform pharmacists about MRPs that occurred outside pharmacy working hours and/or report incidents via the hospital incident reporting systems. In addition, pharmacists may not have identified all MRPs during routine reviews as it is not standard practice for pharmacists to review the laboratory results for all patients daily, or routinely ask patients about possible side-effects (which may have occurred if data were collected under trial conditions).

Other potential limitations include:

- use of a combined outcome event (MRP is an umbrella term for ADRs, ADEs and MEs), which raises two issues:
    - there is a risk that candidate predictors may have opposite predictive effects in components of a combined outcome event, causing their effect to cancel out, leading to important predictors being excluded from the final model[60];
    - differences in the proportion of the outcome components may impact on the predictive accuracy in new datasets[60].

  While use of a combined outcome event is a potential limitation, the predictors were selected based on combined evidence for all outcome components (i.e. MRP subcategories), therefore consistent predictive effects are likely. Potential differences in the proportion of MRP subcategories of the outcome event (MSP MRPs) at the two study sites is reviewed in chapter 8, but this will need to be considered further during MOAT validation studies;

- no data were collected on the frequency or duration of pharmacist reviews of study participants, it is therefore not possible to investigate whether this impacted on MRP data collection;
- retrospective collection of candidate predictor data was used to permit consecutive inclusion of patients from all study wards concurrently (given constraints of time and resources), but this led to a proportion of admissions being excluded due to unavailability of medical records. Despite the consequent reduction in sample size (35 admissions), exclusion is unlikely to bias the analysis as data appears to be MCAR;
- routine clinical records were used as the data source for predictors, resulting in some missing values (where routine assessments not performed / documented);
- there was 'insufficient information' to fully validate eight (0.3%) of the 2,736 MRPs. While this may have biased identification of MSP MRPs (due to measurement error[154]), it is unlikely to have a significant impact on MOAT development given the low number of study admissions affected (five of 1,503, i.e. 0.3%), and likelihood they these were not true MRPs;
- it was necessary to categorise high-risk medicines to permit modelling, but medicines within the same category may not have equivalent risks (e.g. insulin and oral diabetes medication, which are both categorised as 'anti-diabetic medication'). In addition, some medicines are used for more than one indication (e.g. gabapentin can be used for epilepsy and neuropathic pain); the impact of an MRP may therefore be dependent on the indication (e.g. missing doses of gabapentin prescribed for epilepsy may confer higher risk than if used for pain). Similarly, parenteral medicine use was treated as a binary variable, which does not take into account the number or type of parenteral medicine(s) prescribed, or the duration of use. In summary, categorisation is potentially simplistic, albeit necessary to prevent model overfitting;
- when selecting categories for primary diagnosis it was not possible to group 'infections' as a separate category. While this was a logical category from a clinical perspective, as discussed in section 6.2.4.6, it could lead to incorrect categorisation when using the MOAT in clinical practice.

The limitations associated with the use of the Bonferroni correction also need to be considered in terms of the risk of false negative (type II) errors, causing real differences to be overlooked[146]. The Bonferroni correction was used as the primary objective was to test universal null hypotheses, and false positive errors were a greater issue than

false negative errors as the purpose was to establish an answer to these hypotheses, rather than generate hypotheses for further investigation.

Implications

Given the above limitations, it is possible that the present study may underestimate the prevalence of MRPs, although the prevalence was consistent with previous studies[72 78 90]. Subsequent research may be able to improve on the present study by ensuring that all participants receive consistent levels of pharmacist review (for example the duration and frequency of reviews, and consistent use of medicines reconciliation). Use of alternative methods of MRP data collection (for example, use of a dedicated research team) may also improve the consistency of MRP identification, although as discussed in section 6.2.2, prospective identification by pharmacy staff was chosen as it aligns with the proposed purpose of the MOAT, which is to identify patients at risk of MRPs that can be identified during routine clinical practice. Other possible improvements for subsequent research include:

- the use of prospective collection of predictor data, as this may increase the completeness of data, overcoming the need to impute missing values;
- carrying out validation of MRPs as soon as possible after MRP data collection (to permit clarification where needed, overcoming the issue of there being insufficient data to validate all MRPs);
- use of a larger sample to increase the number of outcome events, so permitting the inclusion of more variables in the regression modelling (given the rule of thumb of at least ten 'events per variable'[92]), as this would allow the use of less simplistic categorisation of candidate predictors.

The results presented in this chapter provide support for the potential generalisability of the MOAT. As expected, the comparison of study participants between the two study sites identified statistically significant differences in socioeconomic status and ethnicity, but despite these differences, other participant characteristics were broadly similar, (including age, gender, number of comorbidities, primary diagnoses, number of medicines, and length of hospital stay). This suggests that the MOAT target population may share similar characteristics irrespective of their socioeconomic status and ethnicity. The prevalence of MRPs was also broadly similar between study sites, as were working practices (related to the identification of MRPs during ward duties), suggesting similar medicines optimisation needs, and clinical pharmacy provision between hospitals.

The results also indicated that a number of high-risk medicine categories contained only a small proportion of admissions. For example 'theophylline and aminophylline', and 'lithium', contained only 2.5% and 0.4% respectively. As discussed in section 6.2.4, predictor categories were pre-selected to reduce the risk of selection bias associated with data-driven analysis[62], but on the understanding that it may be necessary to collapse groups if there were insufficient participants to adequately represent the target population (taken by other researchers to be less than 5% of the study sample[42 44 45]). A review of the high-risk medicine categories was therefore required prior to MOAT development.

In addition, the review of missing data (for admissions included in MOAT development) led to the conclusion that the data were likely to be MAR, suggesting that use of multiple imputation is appropriate

6.6 Conclusion

The work presented in this chapter suggests that the observed prevalence of MRPs was consistent with other published research, occurring in 894 (59.5%) of study participants, with 'indication not treated / missing therapy', 'dose too low' and 'dose too high', being the most frequently observed MRP subcategories. There were statistically significant differences in socioeconomic status and ethnicity between study sites, but other participant characteristics were comparable. Where predictor data were missing, it appears to be MAR, permitting the use of multiple imputation to estimate the missing values.

The next chapter will quantify potential variability in MRP identification by pharmacists at the study sites through the development and use of a simulated MRP identification assessment exercise.

Chapter 7: Pharmacists' identification of medication related problems: a validation exercise

7.1 Introduction

Medication related problem (MRP) data for this study were collected by pharmacy staff at the study sites as part of their routine daily clinical assessment of patients. A potential limitation, identified in chapter 6, is the possible impact of knowledge, experience and skills of pharmacists on their ability to identify potential MRPs. The aim of the work presented in this chapter was therefore to quantify potential variability in MRP identification by pharmacists at the study sites. The objectives were to:

- identify the extent of variation among pharmacists in identification of MRPs;
- establish the proportion of potential MRPs identified by pharmacists;
- assess whether identification varies dependant on the type of MRP or potential severity;
- identify implications of variability in MRP identification for the interpretation of overall study findings.

7.2 Methods

This assessment involved the development and then use of a simulated MRP identification assessment exercise, followed by analysis of the results. Each of these stages is described below. The results are described in section 7.3.

7.2.1 Development of the MRP identification assessment exercise

In designing the assessment exercise I considered the need for it to:

- be clinically relevant in terms of current prescribing;
- reflect the breadth and complexity of MRPs commonly experienced by potential study patients;
- be at a level suitable for all study pharmacists irrespective of clinical / service speciality and experience (i.e. aimed at newly qualified, non-specialist pharmacists);
- be simple to complete, both in terms of time taken, and facilities required.

To ensure the assessment met the above requirements I enlisted the assistance of two junior pharmacists working on study wards at Hospital A to advise on design and content. In collaboration, we created four fictitious medication charts (Appendix A7.1), each containing three or four simulated MRPs. The number of medication charts was selected to provide a representative range of patients and MRPs, while being realistic in terms of time taken to review (estimated five to ten minutes to complete the review of all four medication charts). We chose to use paper medication charts to remove the need for access to a computer, and to permit the same assessment process to be used at both study sites. The chart originated from Hospital A, where it was used prior to the introduction of an electronic prescribing system.

7.2.2 Pharmacist completion of MRP identification assessment exercise

All pharmacy staff involved in the study were provided with training prior to commencement of the study (described in section 6.3.3). Following this training, all pharmacists (who may have been required to clinically screen study patients' medication charts) were asked to complete the MRP identification assessment exercise. Prior to undertaking the assessment its purpose was explained, as was the proposed use of the results (i.e. for research purposes to quantify potential variability in MRP identification). The assessment was completed anonymously, as its purpose was not to assess the ability of individual pharmacists. Pharmacists were asked to complete

the assessment exercise individually at a time convenient to them, without discussing the scenarios with other pharmacists. They were advised that the exercise should take between five and ten minutes to complete, but no time limit was set, and pharmacists were not required to record the time taken. Pharmacists were instructed to review each medication chart and list the potential MRPs identified using a form designed for this purpose. They were not told the number of potential MRPs on each medication chart, or the overall number across all four charts. Pharmacists were permitted to refer to references sources if required (to reflect standard clinical practice), but were advised that no other patient information was available (such as medical notes or laboratory results). Once the assessment was completed each pharmacist confirmed return of their form by signing a completion register. The two pharmacists involved in developing the assessment did not take part.

7.2.3 Analysis of results

Each simulated MRP was treated as having a binary outcome in terms of whether or not it was identified by each pharmacist. To identify the extent of variation among pharmacists the median number of simulated MRPs identified per pharmacist and interquartile range (IQR) were calculated (to establish central tendency and variability). The Mann-Whitney U test was also performed to test whether there was a difference between sites (based on a null hypothesis of no difference).

The characteristics of pharmacists (who returned assessment forms) at the two hospitals were compared in terms of their grades and roles. The completion registers were used to obtain pharmacist's names, from which it was possible to establish grade / roles. Grading was based on the Agenda for Change pay scale, which is the National Health Service pay system for employees[155]. Band 6 is the pay point for newly qualified pharmacists, increasing to band 8c for senior / specialist posts[156]. Role was divided into categories based on whether the pharmacists' main role was clinical (i.e. patient facing), or non-clinical, for example service based roles such as technical services or dispensary. Clinical roles were further divided into those based predominantly on study wards, and those predominantly based in other clinical areas (such as surgery and paediatrics). A similar comparison was also made between pharmacists who did, or did not, return an assessment form. Fisher's exact tests were performed to test for differences between groups (chosen due to size of the sub-groups and *expected* distribution of results in the absence of association).

The number / proportion of pharmacists who identified each MRP was calculated to give the percentage agreement for each MRP, and the results reported separately for each MRP by study site, in addition to combined results for both sites. To test if there was evidence for a difference in the proportion of simulated MRPs identified by pharmacists at the two study sites a chi-square test was performed. Sub-group analysis was performed to identify potential differences in percentage agreement for individual MRPs between sites (using Fisher's exact tests).

Randolph's kappa was used to assess chance-adjusted agreement between pharmacists (calculated using Randolph's online kappa calculator[157]). Randolph's kappa was chosen as it is a multiple rater kappa test. Other multi-rater tests are available, including Light's, Hubert's, and Fleiss' kappa[158], but these are 'fixed-marginal' versions, unlike Randolph's kappa, which is 'free-marginal'[159]. Fixed-marginal tests are recommended where the marginals (i.e. the proportion that should be distributed into each category) is known to raters prior to the assessment, whereas free-marginal versions are recommended when raters are not restricted in the number of cases assigned to each category[159], as was the case for the MRP assessment exercise. When fixed-marginal varieties of kappa are used in free-marginal assessments the value of kappa can vary significantly dependent on the marginal distributions, a finding attributed to the prevalence (proportion of cases), and bias (defined as the bias of one rater relative to another)[160]. Kappa values were interpreted as poor if lower than 0.00, slight if 0.00 to 0.20, fair if 0.21 to 0.40, moderate if 0.41 to 0.60, substantial if 0.61 to 0.80, and almost perfect if between 0.81 and 1.00[161 162].

Exploratory analysis was performed to assess whether identification was influenced by either the type (MRP subcategory) or potential severity of the simulated MRPs. Basger's aggregated classification system[114] was used to classify the MRPs, and severity was assessed using a validated visual analogue scale for medication errors[163] (as used for the MRP study data, described in chapters 6 and 8 respectively). For the severity rating I categorised MRPs as either potentially 'minor', or 'moderate or severe'. Where possible this assessment was based on ratings assigned to similar MRPs in the study dataset. The percentage agreement for each classification and severity category were calculated, and chi-square tests performed to assess evidence for statistically significant differences.

Where appropriate, the Bonferroni correction was applied to the probability (*p*) values to account for the risk of type I (false positive) errors associated with multiple

analyses[146]; Bonferroni corrected *p* values were calculated based on the number of comparisons (to maintain the critical *p* level over all tests at 0.05).

7.3 Results

The MRP identification assessment comprised four medication charts, with a total of 13 simulated MRPs. Fifty nine pharmacists completed the assessment for all four medication charts and were included in the analysis (30 from Hospital A and 29 from Hospital B). This represented all pharmacists involved in the study at Hospital A (excluding the two pharmacists involved in developing the assessment), and 66% of the 44 pharmacists at Hospital B. One further pharmacist at Hospital B returned an assessment form, but it was not possible to include their results in the analysis as only three of the four medication charts had been assessed (the number of raters needs to be consistent for all assessments when calculating the kappa statistic). Due to anonymisation of the forms it was not possible to identify the pharmacist concerned to find out why their assessment was incomplete.

The grade and main role of pharmacists who completed the assessment is summarised in Table 26. This shows that the highest proportion of pharmacists at Hospital A were graded 8a (46.7%), whereas the highest proportion at Hospital B were graded at band 7 (43.3%), although there was no evidence for a statistically significant difference in pharmacists' grades between the two sites ($p = 0.381$). A difference was also observed between the main roles of pharmacists at the two sites, with a higher proportion of pharmacists at Hospital B working predominantly on study wards (73.3% compared to 43.3% at Hospital A). Given the Bonferroni corrected p value of 0.025 the difference was not statistically significant ($p = 0.051$), although the lack of evidence against the null hypothesis may be explained by the small sample size. The apparent differences between sites in terms of the band and role of pharmacists may be explained by the greater number of study wards at Hospital B compared to Hospital A (19 and 11 respectively, described in section 6.3.2), leading to Hospital B requiring proportionally more junior pharmacists (bands 6 or 7) to provide routine pharmacy services to study wards.

Table 26 – Grade and role of pharmacists who returned the assessment

Grade / main role	Assessment form returned	
	Hospital A (n = 30) n (%)	Hospital B (n = 30†) n (%)
Grade*		
Band 6	7 (23.3)	5 (16.7)
Band 7	6 (10.0)	13 (43.3)
Band 8a	14 (46.7)	9 (30.0)
Band 8b	2 (6.7)	2 (6.7)
Band 8c	1 (3.3)	1 (3.3)
Main role		
Clinical role (predominantly working on study wards)	13 (43.3)	22 (73.3)
Clinical role (predominantly working on non-study wards e.g. surgery, paediatrics)	7 (23.3)	2 (6.7)
Non-clinical role (e.g. technical services, dispensary)	10 (33.3)	6 (20.0)

* Grade based on Agenda for Change pay scale

† All returned assessment forms, therefore includes the incomplete form

Bonferroni adjusted *p* value used to judge statistical significance 0.025 (based on 2 statistical tests)

Table 27 summarises the response rates at Hospital B by the grade and role of pharmacists, permitting the potential impact of the completion rate of 66% to be examined. Interpretation is limited by the small sample size, but results suggest that response rates were relatively consistent for all grades of pharmacists (64.3% to 71.4% for bands 6 to 8b), with no evidence for a statistically significant difference ($p = 0.97$). There appears to be a difference in response rate related to pharmacists' role, with forms returned by 75.9% of pharmacists working on study wards compared to 40.0% and 60.0% for the other categories. This difference was not statistically significant ($p = 0.193$), but may have contributed to the increased variability observed in roles at Hospital B compared to Hospital A (Table 26).

Table 27 – Response rates for pharmacists at Hospital B

Grade / main role	Number of pharmacists (n = 44†)	Assessment returned (n = 30†) n (% response rate)
Grade*		
Band 6	7	5 (71.4)
Band 7	19	13 (68.4)
Band 8a	14	9 (64.3)
Band 8b	3	2 (66.7)
Band 8c	1	1 (100.0)
Main role		
Clinical role (predominantly working on study wards)	29	22 (75.9)
Clinical role (predominantly working on non-study wards e.g. surgery, paediatrics)	5	2 (40.0)
Non-clinical role (e.g. technical services, dispensary)	10	6 (60.0)

* Grade based on Agenda for Change pay scale

† All returned assessment forms, therefore includes the incomplete form

Bonferroni adjusted *p* value used to judge statistical significance 0.025 (based on 2 statistical tests)

7.3.1 Variation among pharmacists in identification of medication related problems

As summarised in Table 28, each pharmacist identified between five and 13 simulated MRPs, with a median of 12 identified per pharmacist across both study sites (IQR ten to 12).

Pharmacists at Hospital A identified a median of 12 simulated MRPs, compared to 11 at Hospital B, with a greater proportion of pharmacists identifying all 13 MRPs at Hospital A (30.0% compared to 13.3% at Hospital B). Greater variability was observed at Hospital A, shown by the range in the number of MRPs identified (from five to 13 at Hospital A, and 7 to 13 at Hospital B), and IQR (ten to 13 at Hospital A, compared to ten to 12 at Hospital B). Despite these differences there was no evidence for a statistically significant difference between study sites ($p = 0.357$), but this may be due to the small sample size.

Table 28 – Number of simulated medication related problems identified per pharmacist

Number of simulated medication related problems identified per pharmacist	Frequency Hospital A (pharmacists = 30) n (% of pharmacists)	Frequency Hospital B (pharmacists = 29) n (% of pharmacists)	Combined frequency (pharmacists = 59) n (% of pharmacists)
5	1 (3.3)	0	1 (1.7)
7	1 (3.3)	1 (3.3)	2 (3.4)
8	2 (6.7)	0	2 (3.4)
9	2 (6.7)	5 (16.7)	7 (11.9)
10	4 (13.3)	6 (20.0)	10 (16.9)
11	3 (10.0)	4 (13.3)	7 (11.9)
12	8 (26.7)	9 (30.0)	17 (28.8)
13	9 (30.0)	4 (13.3)	13 (22.0)
Median	**12**	**11**	**12**
Interquartile range	**10-13**	**10-12**	**10-12**

7.3.2 Proportion of potential medication related problems identified

A brief description of each simulated MRP and the number / percentage of pharmacists who identified each MRP (percentage agreement) is given in Table 29 (results shown separately for each study site and also as a combined result for both sites). The overall percentage agreement at Hospitals A and B was 85% and 84% respectively, with no evidence for a difference between sites (p = 0.617). Sub-group analysis also found no evidence for differences in percentage agreement for individual MRPs between sites (using Bonferroni corrected p value of 0.0036), therefore the remaining analyses were performed using the combined results.

The combined percentage agreement for each simulated MRP ranged from 70% to 98%, with an overall agreement for all MRPs of 84.5%. Randolph's kappa was calculated to assess the chance-adjusted agreement for each MRP, with results ranging from 0.14 to 0.93, suggesting agreement varied from 'slight' to 'almost perfect'[161 162]. Randolph's kappa for the combined results (both study sites and all MRPs) was 0.50, suggesting moderate agreement[161 162 164].

Table 29 – Medication related problem identification and chance-adjusted agreement

Medication chart	Description of simulated medication related problem (MRP)	MRP identified Hospital A (pharmacists = 30) n (% of pharmacists)	MRP identified Hospital B (pharmacists = 29) n (% of pharmacists)	Combined results (pharmacists = 59) n (% of pharmacists)	Kappa (combined results)	*p* value
1	Methotrexate and trimethoprim prescribed concurrently	21 (70.0)	22 (75.9)	43 (72.9)	0.20	0.771 (Fisher's exact)
	Paracetamol prescribed regularly (maximum dose) plus when required	29 (96.7)	29 (100.0)	58 (98.3)	0.93	1.000 (Fisher's exact)
	Methotrexate prescribed daily (not weekly)	30 (100.0)	27 (93.1)	57 (96.6)	0.87	0.237 (Fisher's exact)
2	Furosemide prescribed at 6pm (no indication that patient was catheterised)	25 (83.3)	21 (72.4)	46 (78.0)	0.30	0.360 (Fisher's exact)
	Two loop diuretics (furosemide and bumetanide) co-prescribed	29 (96.7)	27 (93.1)	56 (95.0)	0.80	0.612 (Fisher's exact)
	Nicorandil prescribed once daily (usually twice daily)	24 (80.0)	25 (86.2)	49 (83.1)	0.43	0.731 (Fisher's exact)
3	Simvastatin prescribed in the morning (usually taken in evening)	20 (66.6)	21 (72.4)	41 (69.5)	0.14	0.779 (Fisher's exact)
	Digoxin dose prescribed as milligrams (rather than micrograms)	26 (86.7)	20 (69.0)	46 (78.0)	0.30	0.125 (Fisher's exact)
	Apixaban prescribed once daily (usually twice daily)	24 (80.0)	28 (96.6)	52 (88.1)	0.58	0.103 (Fisher's exact)
	Simvastatin 40mg daily co-prescribed with amlodipine (maximum dose of simvastatin is 20mg when co-prescribed)	25 (83.3)	21 (72.4)	46 (78.0)	0.30	0.360 (Fisher's exact)
4	Patient prescribed a penicillin containing antibiotic (when allergic to penicillin)	24 (80.0)	27 (93.1)	51 (86.4)	0.52	0.254 (Fisher's exact)
	Patient not received phenytoin for 3 days as nil-by-mouth (no alternative formulation prescribed)	28 (93.3)	20 (69.0)	48 (81.4)	0.38	0.021 (Fisher's exact)
	Prophylactic enoxaparin co-prescribed with therapeutic tinzaparin (therapeutic duplication)	27 (90.0)	28 (96.6)	55 (93.2)	0.74	0.612 (Fisher's exact)
Overall results (all MRPs)		**332 (85.1)**	**316 (83.8)**	**648 (84.5)**	**0.50**	**0.617** (Chi-square)

Bonferroni adjusted *p* value used to judge statistical significance 0.0036 (based on 14 statistical tests)

7.3.3 Analysis by MRP subcategory

Categorisation of the 13 simulated MRPs (using Basger's aggregated classification system) resulted in seven subcategorises (Table 30). Percentage agreement for the seven subcategories ranged from 73% to 94%, with strong evidence for a difference among subcategories ($p < 0.0001$).

Table 30 – Medication related problem (MRP) identification analysed by MRP subcategory

Medication related problem (MRP) subcategory*	Description of simulated MRP	Percentage agreement	Percentage agreement (grouped by MRP subcategory)
1.1 Inappropriate drug	Patient prescribed a penicillin containing antibiotic (when allergic to penicillin)	86%	86%
1.2 No indication for drug / duplication	Two loop diuretics (furosemide and bumetanide) co-prescribed	95%	94%
	Prophylactic enoxaparin co-prescribed with therapeutic tinzaparin (therapeutic duplication)	93%	
1.3 Interaction	Methotrexate and trimethoprim prescribed concurrently	73%	75%
	Simvastatin 40mg daily co-prescribed with amlodipine (maximum dose of simvastatin is 20mg when co-prescribed)	78%	
3.1 Drug dose too low	Nicorandil prescribed once daily (usually twice daily)	83%	86%
	Apixaban prescribed once daily (usually twice daily)	88%	
3.2 Drug dose too high	Paracetamol prescribed regularly (maximum dose) plus when required	98%	94%
	Methotrexate prescribed daily (not weekly)	97%	
	Digoxin dose prescribed as milligrams (rather than micrograms)	78%	
5.1 Inappropriate timing of administration / dosing by prescriber	Furosemide prescribed at 6pm (no indication that patient was catheterised)	78%	73%
	Simvastatin prescribed in the morning (usually taken in evening)	70%	
5.4 Drug not taken / administered at all	Patient not received phenytoin for 3 days as nil-by-mouth (no alternative formulation prescribed)	81%	81%

* Classified using Basger's aggregated classification system[114]

7.3.4 Analysis by severity

Table 31 summarises the percentage agreement for MRPs based on their potential severity. Although agreement was higher for the MRPs rated as 'moderate or severe' compared to those rated as 'minor' (86% and 81% respectively), there was no evidence for a statistically significant difference (p = 0.122).

Table 31 – Medication related problem identification analysed by potential severity

Potential severity of medication related problem (MRP)*	Description of simulated MRP	Percentage agreement	Percentage agreement (grouped by severity)
Moderate or severe outcome	Methotrexate and trimethoprim prescribed concurrently	73%	86%
	Paracetamol prescribed regularly (maximum dose) plus when required	98%	
	Methotrexate prescribed daily (not weekly)	97%	
	Nicorandil prescribed once daily (usually twice daily)	83%	
	Digoxin dose prescribed as milligrams (rather than micrograms)	78%	
	Apixaban prescribed once daily (usually twice daily)	88%	
	Simvastatin 40mg daily co-prescribed with amlodipine (maximum dose of simvastatin is 20mg when co-prescribed)	78%	
	Patient prescribed a penicillin containing antibiotic (when allergic to penicillin)	86%	
	Patient not received phenytoin for 3 days as nil-by-mouth (no alternative formulation prescribed)	81%	
	Prophylactic enoxaparin co-prescribed with therapeutic tinzaparin (therapeutic duplication)	93%	
Minor outcome	Furosemide prescribed at 6pm (no indication that patient was catheterised)	78%	81%
	Two loop diuretics (furosemide and bumetanide) co-prescribed	95%	
	Simvastatin prescribed in the morning (usually taken in evening)	70%	

* Classified using a validated visual analogue scale for medication errors[163]

7.4 Discussion

Key findings

The assessment found there were no statistically significant differences between results from the two study sites. Each study pharmacist identified between five and 13 of the simulated MRPs (out of a total of 13), with a median of 12 and IQR of ten to 12. The overall percentage agreement (proportion of MRPs identified) was 84.5%, with a kappa coefficient of 0.50. No statistically significant differences were found between the percentage agreement for MRPs graded as potentially 'minor' compared to 'moderate or severe', but a difference was found among different types of MRPs (MRP subcategories).

Interpretation

While no statistically significant differences were seen in the characteristics of pharmacists between study sites, or between pharmacists who completed the assessment compared to those who did not, this may be due to the limited sample size. Despite not reaching statistical significance, one could hypothesise that the differences observed between sites (in grading and role) may explain other variability in the results of the assessment. For example it is possible that the slightly lower variability in the number of MRPs identified per pharmacist at Hospital B compared to Hospital A may be explained by the higher proportion of pharmacists at Hospital B who routinely worked on study wards, as they would be more familiar with the scenarios used for the MRP identification exercise. Similarly, the higher proportion of senior pharmacists (band 8a) at Hospital A may explain the higher median number of MRPs identified, due to increased experience and knowledge. A larger study, or more detailed analysis (incorporating the grade and role of individual pharmacists), would be required to investigate these hypotheses further.

Regarding variation in the proportion of MRPs identified, there is no standardised interpretation for percentage agreement or kappa statistic, with the decision on what level is acceptable being determined by the purpose of the assessment[161]. In health research 80% is often recommended as the minimum acceptable percentage agreement[164], with a kappa of 0.41 to 0.60 interpreted as demonstrating 'moderate agreement'[161 162].

While the overall percentage agreement exceeded 80%, agreement for individual MRPs ranged from 69.5% to 98.3%, with six of the 13 MRPs achieving a percentage agreement lower than 80% (at one or both study sites), causing the corresponding

reduction in the overall kappa statistic. These MRPs, together with possible explanations for the variability, are listed below.

1. **Medication chart 1, methotrexate and trimethoprim prescribed concurrently (percentage agreement 72.9%, kappa 0.20)**
   This is potentially serious interaction[165], therefore the result cannot be explained by differing professional judgement. Although it is possible that pharmacists were less vigilant when completing the assessment (compared to routine practice) other MRPs had percentage agreement close to 100%, suggesting this variability reflects differences in pharmacists' awareness of the interaction.
2. **Medication chart 2, furosemide prescribed at 6pm with no indication of catheterisation (percentage agreement 78.0%, kappa 0.30)**
   Furosemide is a loop diuretic, therefore should ideally be administered no later than early afternoon to prevent nocturia (excessive urination at night)[165]. As identification of MRPs is influenced by professional judgement, some pharmacists may not have considered this MRP to be clinically relevant, particularly as the patient was hospitalised therefore able to request assistance with toileting overnight.
3. **Medication chart 3, simvastatin prescribed in the morning rather than evening (percentage agreement 69.5%, kappa 0.14)**
   This may be another example of differing professional judgement regarding clinical relevance. While there is evidence that simvastatin is best taken in the evening, this is not true for all statins[166], there is also recognition that adherence may be compromised by multiple dosing[167]. Both factors may have influenced pharmacists' decisions regarding whether this was a potential MRP.
4. **Medication chart 3, digoxin prescribed as milligrams rather than micrograms (percentage agreement 78.0%, kappa 0.30)**
   Although this is a potentially serious overdose, the likelihood of the error reaching the patient is minimal as it would involve administering 1,000 times the standard dose. This may cause pharmacists to be less alert to this type of error.
5. **Medication chart 3, interaction between simvastatin and amlodipine (percentage agreement 78.0%, kappa 0.30)**
   Identification of this MRP may have been influenced by differences in medicines usage between the study sites. Amlodipine was commonly used at Hospital A, whereas felodipine was the more commonly used calcium channel blocker at Hospital B. This may have resulted in lower awareness of the interaction at Hospital B, leading to reduced identification (72.4% compared to 83.3% at Hospital A).

6. **Medication chart 4, phenytoin not administered for 3 days (percentage agreement 81.4%, kappa 0.38)**

   The difference in percentage agreement between sites (90% at Hospital A and 69% at Hospital B) may be due to pharmacists at Hospital B being unfamiliar with the medication charts used for the assessment (the chart used for the assessment originated at Hospital A, where different 'codes' were used to record non-administration). While the difference in percentage agreement between the two sites did not reach statistical significance (once Bonferroni correction applied), there was some evidence for a difference (p = 0.021).

The analysis of MRPs by MRP subcategories found a statistically significance difference between categories, with the results suggesting that 'interactions' and 'inappropriate timing of administration' may be less likely to be identified by pharmacists than the other types of MRPs. This is consistent with the findings (discussed above) regarding the impact of professional judgement (in relation to inappropriate timing of administration), and pharmacist's awareness of the two interactions included in the assessment. Caution is required when interpreting these results though, due to:

- the relatively small number of MRPs (and MRP subcategories) included in the assessment;
- the variation in percentage agreement among MRPs within subcategories (for example percentage agreement for the MRPs classified as 'dose too high' ranged from 78% to 98%);
- two subcategories consisting of only one MRP, which may not have been representative of other MRPs within the category (in terms of percentage agreement).

The analysis of MRP identification by severity found no evidence for a difference between MRPs categorised as 'minor' compared to 'moderate or severe', but as above, caution is required due to the small number of MRPs (with only three categorised as potentially minor).

In summary, the results of this assessment suggest acceptable agreement among pharmacists in relation to MRP identification. It is possible that some of the observed variability in MRP identification was related to limitations in the assessment method, but additional factors linked to the knowledge, experience and skills of study pharmacists (such as awareness of interactions, differences in professional judgement,

and perceived likelihood of the error reaching the patient) also appear to have an impact. MRP identification also appears to be associated with MRP classification subcategory, but not with MRP severity, although these findings need to be interpreted with caution.

Strengths and limitations

A strength of the research presented in this chapter was its reproducibility, between individual pharmacists and study sites, which would have been difficult to achieve using non-standardised assessment methods, such as observation. Other strengths include the input of practising pharmacists (working with study patients) to inform development of the assessment exercise, and the time-efficient nature of the assessment, from both the assessor and assessee perspectives.

Potential limitations include:

- the choice of medication chart used for the assessment (given that Hospital A had recently introduced an electronic prescribing system, and Hospital B used a slightly different paper chart). This represented a compromise between developing an assessment that was true to practice, and one that could be used across both study sites, and was time efficient and simple to complete;
- the inability to assess all types of MRPs using this method, for example, it was not possible to assess pharmacists' ability to identify medicines reconciliation discrepancies, the need for dosage adjustments due to organ dysfunction, and administration or monitoring errors;
- the response rate at Hospital B (66%), as this is a potential source of selection bias, and prevented a direct comparison between study pharmacists at the sites;
- the use of anonymisation, meaning it was not possible to trace the pharmacist who submitted the incomplete form, or carry out more detailed analyses based on grade, role, and/or experience;
- the size of the study, limiting the statistical inferences that could be made;
- the length of the assessment (number of medication charts / simulated MRPs), as although the length of the assessment was chosen to limit the amount of time needed (to enhance completion) a longer assessment would have permitted:
    - inclusion of a greater number and/or a broader range of MRPs (to permit a more robust comparison between MRP classification and severity categories);
    - inclusion of medication charts with no MRPs (to more accurately reflect practice).

As in chapter 6, the limitations associated with the use of the Bonferroni correction also need to be considered[146]. The Bonferroni correction was used as the primary objective was to test universal null hypotheses (that identification was unrelated to study site, MRP classification and severity). In addition, false positive errors were a greater issue than false negative errors as the purpose was to establish an answer to this hypothesis, rather than generate hypotheses for further investigation.

Implications

Based on this assessment, the best estimate is that pharmacists are likely to identify approximately 85% of study MRPs (although this assumes similar percentage agreements for MRP subcategories not included in the assessment). While a percentage agreement of 80% may be considered acceptable in health research[164], identification of less than 100% of outcome events may be subject to criticism. This potential variability therefore needs to be recognised as a limitation of the MRP data collection method, although as discussed in section 6.2.2, prospective identification by pharmacy staff was chosen as it aligns with the proposed purpose of the Medicines Optimisation Assessment Tool (MOAT^{TM}), which is to identify patients at risk of MRPs that can be identified during routine clinical practice. The potential variability in identification also needs to be taken into account when interpreting study data, for example the observed prevalence of MRPs is likely to be underestimated. As the variability appears to be unrelated to MRP severity it may also impact on identification of patients with the outcome event (i.e. at least one moderate or severe preventable MRP). This has the potential to reduce the accuracy and replicability of MOAT predictions, which highlights the need for robust external validation of the MOAT (following initial development), including the possible need for updating or recalibratation[94].

7.5 Conclusion

The work presented in this chapter suggests that pharmacists at the study sites are likely to identify approximately 85% of study MRPs. Identification appears to be unaffected by MRP severity, but may be associated with the type of MRP (classification subcategory). This variability in MRP identification needs to be recognised as a limitation, and the implications will need to be considered when interpreting overall study findings.

The next chapter will describe the severity assessment of MRPs identified by pharmacists at the study sites. This was undertaken to identify patients with the outcome event, namely at least one moderate or severe preventable MRP.

Chapter 8: Analysis of outcome events

8.1 Introduction

As discussed in chapter 6, previous research suggests that a significant proportion of hospitalised patients experience medication related problems (MRPs), with estimates as high as 81%[78]. A prognostic model developed to predict MRPs would therefore lead to a high proportion of patients being labelled as high-risk, potentially leading to inefficient workload management. I therefore chose to select 'clinically significant' MRPs as the outcome event for the Medicines Optimisation Assessment Tool (MOAT^{TM}), based on potential severity. Similarly, I chose to use 'preventable' MRPs, to enable the MOAT to identify patients with MRPs that are amenable to pharmacist intervention. The outcome event selected for development of the MOAT was therefore moderate or severe preventable MRPs (MSP MRPs). MRP preventability was assessed at the point of identification (see section 6.3.3), but it was necessary to select a suitable method to rate the potential severity of MRPs, which is described in the present chapter.

The aim of the MOAT is to permit targeting of hospital patients most in need of pharmacists' input. This will require accurate prediction of those at risk, but may also require understanding of 'when' patients are most at risk, as it is possible that risk will vary throughout an admission to hospital. An awareness of the stage of patient stay when patients are most likely to experience MSP MRPs may therefore be useful in informing implementation of the MOAT. Similarly, an understanding of working practices at the study sites, for example the proportion of MSP MRPs identified by pharmacists working at ward level compared to those identified in the pharmacy department, may permit future MOAT users to assess the applicability of the MOAT to their own setting.

In addition, a potential limitation of the present study, as discussed in chapter 6, was use of a combined outcome measure; MRP is an umbrella term for a number of 'outcome components' (i.e. MRP subcategories), therefore significant differences in the proportion of these components in new datasets could impact on the MOAT's predictive accuracy[60]. To permit comparison with future validation datasets it was therefore necessary to determine the proportion of each 'outcome component' in the developmental dataset, that is to say, to establish the proportion of each MRP subcategory for the MOAT's outcome event (MSP MRPs). As two hospitals were used for MOAT development, it was also possible to compare the proportion of each MRP

subcategory between sites, so providing an indication of potential generalisability of the MOAT. That is to say, consistency in the composition of MSP MRPs between hospital sites may suggest greater generalisability.

The aim of the work presented in this chapter was therefore to describe the MRP severity assessment, and descriptive analyses of the outcome event, namely MSP MRPs. The objectives were to:

- identify MSP MRPs (the outcome event for development of the MOAT);
- describe descriptive information on the MSP MRPs (to inform future implementation of the MOAT);
- determine the composition of the 'outcome components' (i.e. MRP subcategories) of the MSP MRPs to permit comparison with future validation datasets, and provide some indication of potential generalisability of the MOAT.

8.2 Methods

The methods are presented in three sections: the severity rating of MRPs, descriptive analysis of MSP MRPs, and the prevalence of MSP MRPs in the study sample. Each is described below, with the results described in section 8.3.

8.2.1 Severity rating of MRPs

As no established grading for MRPs is available, severity was classified using a validated visual analogue scale developed originally for medication errors[163], used previously to severity rate MRPs by Rashed *et al*[168]. This method requires four experienced healthcare professionals (pharmacists, medical, or nursing staff) to independently score each event in terms of potential patient outcomes on a scale of zero to ten. Zero represents a case with no potential adverse effect on the patient, and ten a case that would result in death. The mean score for each event is then used as an index of severity, with a score of less than three being considered as a minor outcome (very unlikely to have an adverse effect), a score of three to seven considered as moderate (likely to cause some adverse effects or interfere with therapeutic goals, but very unlikely to result in death or lasting impairment), and a score of greater than seven considered to be a severe outcome (likely to cause death or lasting impairment). Dean *et al*[163] suggest that at least one judge be selected from each of the three professions to facilitate ownership of the results[163]; I therefore chose a hospital pharmacist, senior nurse, and consultant physician (plus myself) to form the MOAT

expert panel. All had medication / patient safety experience, and had worked in clinical practice for between 12 and 30 years.

These expert panel members validated each MRP, that is to say they decided whether it was a 'true MRP' (described in chapter 6, section 6.3.3). Validated MRPs that were considered to be preventable (preventability assessment also described in section 6.3.3), were then assessed for severity by the expert panel as described above. To facilitate the assessment process, and aid consistency, expert panel members met twice prior to carrying out the MRP assessment. During the first meeting we discussed the proposed assessment method and general administrative issues, then during the second meeting we discussed 35 randomly selected MRPs in terms of their validation and potential severity (assessed / graded individually by panel members in advance of the meeting). The outcome of this 'sample grading' was shared with panel members (for reference purposes), and the following conventions were agreed:

- MRPs were graded on the basis of the potential severity *if not identified during the hospital stay*;
- severity ratings were based on perceived risk (rather than local policy / guidance). Similarly, national guidance was taken into account if applicable;
- MRPs were graded based on the potential severity to the patient involved, rather than wider implications. For example, when reviewing MRPs related to unnecessary antibiotic use, panel members were asked to consider the direct risk to the patient, such as toxicity, rather than wider societal concerns, such as antimicrobial resistance.

8.2.2 Descriptive analysis of MSP MRPs

MSP MRP data were analysed to establish:

- whether the MSP MRPs were identified during ward visits or in the pharmacy department;
- whether MSP MRPs were resolved by pharmacy staff or other healthcare professionals;
- the stage in 'patient stay' when MSP MRPs identified, classified as during / before first ward review by pharmacist, during the remainder of the inpatient stay, or during clinical screening at discharge;
- whether MSP MRPs were medicines reconciliation discrepancies.

This was to permit comparison between study sites in terms of working practices, inform MOAT implementation, and permit a comparison with equivalent data collected for 'all MRPs' (i.e. all validated MRPs irrespective of severity and preventability, as reported in chapter 6).

MSP MRPs were also categorised using Basger's MRP classification system[114]. This was to permit comparison of the subcategories:

- between study sites (to provide some indication of potential generalisability of the MOAT);
- with 'all MRPs' (to identify potential differences);
- with future validation datasets (as significant differences in the proportion of MRP subcategories in new datasets could impact on the MOAT's predictive accuracy[60]).

Statistical analysis was performed to test for differences between study sites. Chi-square or Fisher's exact tests were used, with the choice based on the *expected* distribution of results in the absence of association, with Fisher's exact test used when an expected cell frequency was less than five. The Bonferroni correction was applied to the probability (*p*) values to account for the risk of type I (false positive) errors associated with multiple analyses[146]; Bonferroni corrected *p* values were calculated based on the number of comparisons (to maintain the critical *p* level over all tests at 0.05).

8.2.3 Prevalence of outcome event

The prevalence of the outcome event (i.e. the proportion of study admissions that experienced at least one MSP MRP), and median number MSP MRPs per admission (of admissions with an outcome event) were calculated, and statistical analysis performed to test for differences between study sites.

8.3 Results

The results are presented in three sections: severity rating of MRPs, descriptive analysis of MSP MRPs, and the prevalence of MSP MRPs in the study sample. Each will be described in turn.

8.3.1 Severity rating of MRPs

As discussed in section 6.4.5.1, a total of 2,614 MRPs were judged by the expert panel to be true MRPs. As shown in Table 32, five were non-preventable (all adverse drug reactions), and 1,456 were rated as 'minor severity'. The remaining 1,153 MRPs were therefore MSP MRPs (i.e. the outcome event selected for development of the MOAT). There was no evidence for a statistically significant difference in the proportion of MRPs rated as 'minor severity' compared to 'moderate or severe' between study sites (p = 0.445).

Table 32 – Identification of moderate or severe preventable medication related problems

	Medication related problems (MRPs)		
	Hospital A n (% of MRPs)	**Hospital B** n (% of MRPs)	**All patients** n (% of MRPs)
Number of validated MRPs (as reported in chapter 6)	1,697 (100)	917 (100)	2,614 (100)
Number of non-preventable MRPs (therefore excluded as outcome events)	5 (0.3)	0 (0)	5 (0.2)
Number of preventable MRPs rated as 'minor severity' (therefore excluded as outcome events)	935 (55.1)	521 (56.8)	1,456 (55.7)
Total MRPs remaining (i.e. moderate or severe preventable MRPs)	**757 (44.6)**	**396 (43.2)**	**1,153 (44.1)**

Of the 1,153 MSP MSPs, only one (from Hospital A) was rated as potentially 'severe'. This involved a patient prescribed prophylactic heparin (for thromboprophylaxis) despite having a raised International Normalised Ratio (of 7.9), increasing the risk of bleeding[165]. The remaining MSP MRPs were all graded as 'moderately severe'. The median scores and interquartile ranges (IQR) for 'moderately severe' MRPs were the same for both study sites (median 3.25, IQR 3.0 to 3.75); the median was calculated (rather than the mean) due to the non-parametric distribution of data.

8.3.2 Descriptive analysis of MSP MRPs

The descriptive analysis results are summarised in Table 33. This shows that the majority of MSP MRPs (93.6%) were identified during routine ward visits, resolved by pharmacy staff (98.2%), and identified during (or before) the first review of the patient (73.9%). Given the Bonferroni corrected *p* value of 0.0125, there were no statistically significant differences between study sites, although there was weak evidence for a difference in the proportion of medicines reconciliation discrepancies, accounting for 55% at Hospital A, and 47.5% Hospital B ($p = 0.016$).

Table 33 – Descriptive data for moderate or severe preventable medication related problems

	Hospital A (MSP MRPs = 757) n (% of MSP MRPs)	**Hospital B (MSP MRPs = 396)** n (% of MSP MRPs)	**All patients (MSP MRPs = 1,153)** n (% of MSP MRPs)	***p* value (test for difference between study sites)**
When MSP MRP identified:				
During ward visit	709 (93.7)	370 (93.4)	1,079 (93.6)	0.812 (Chi-square)
In the pharmacy department	43 (5.7)	22 (5.6)	65 (5.6)	
Other (pharmacist referral / reported via hospital incident reporting system)	5 (0.7)	4 (1.0)	9 (0.8)	
Who resolved MSP MRP:				
Pharmacy staff	740 (97.8)	392 (99.0)	1,132 (98.2)	0.136 (Chi-square)
Other healthcare professionals	17 (2.2)	4 (1.0)	21 (1.8)	
Stage in patient stay MSP MRP identified:				
During first ward review (or before)	559 (73.8)	293 (74.0)	852 (73.9)	0.934 (Chi-square)
Remainder of inpatient stay	118 (15.6)	58 (14.6)	176 (15.3)	
Clinical screening at discharge	80 (10.6)	42 (10.6)	122 (10.6)	
Missing data	0	3 (0.8)	3 (0.3)	
Medicines reconciliation discrepancy	416 (55.0)	188 (47.5)	604 (52.4)	0.016 (Chi-square)

MSP MRP = moderate or severe preventable medication related problem

Bonferroni adjusted *p* value used to judge statistical significance 0.0125 (based on 4 statistical tests)

The overall results (combining MSP MRPs from both study sites) are comparable with the descriptive analysis of 'all MRPs' (Table 22, chapter 6), where 93.2% of MRPs were identified during routine ward visits, 97.9% resolved by pharmacy staff, 74.6% identified during (or before) the first review of the patient, and 55.9% were medicines reconciliation discrepancies. While no statistically significant differences were found between study sites in the descriptive analysis of MSP MRPs, the analysis of 'all MRPs' found evidence for a difference between the two study sites in the stage of the

patient stay when MRPs were identified. As discussed in chapter 6, this may have been due to a greater proportion of pharmacists' time being spent reviewing new rather than existing patients at Hospital B (linked to undertaking medicines reconciliation). The lack of evidence for a difference between sites for MSP MRPs may therefore suggest that the occurrence of MSP MRPs is less sensitive to differences in hospital processes (such as the proportion of patients receiving medicines reconciliation) compared to 'all MRPs'.

Table 34 gives the breakdown of MSP MRPs identified during clinical screening at discharge, in terms of whether they occurred prior to the discharge prescription being written (i.e. those that had occurred, but not identified / resolved, such as medicines reconciliation omissions), or when the discharge prescription was written (e.g. transcription errors). This shows that the majority of MSP MRPs occurred when discharge prescriptions were written (66.4%), with no statistically significant difference between study sites (p = 0.394).

Table 34 – Breakdown of identification of moderate severe preventable medication related problems when discharge prescription screened

When MSP MRP occurred	**Hospital A (MSP MRPs at discharge = 80)** n (% of MSP MRPs at discharge)	**Hospital B (MSP MRPs at discharge = 42)** n (% of MSP MRPs at discharge)	**All patients (MSP MRPs at discharge = 122)** n (% of MSP MRPs at discharge)	***p* value (test for difference between study sites)**
Prior to writing discharge prescription	29 (36.2)	12 (28.6)	41 (33.6)	0.394 (Chi-square)
When discharge prescription written	51 (63.8)	30 (71.4)	81 (66.4)	

MSP MRP = moderate or severe preventable medication related problem

This is comparable with the results for 'all MRPs' (Table 23, chapter 6) where 70.4% of MRPs occurred when discharge prescriptions were written, with no statistically significant difference between study sites.

The classification of the 1,153 MSP MRPs is summarised in Table 35. This shows that the most frequently identified subcategory was 'indication not treated / missing therapy', accounting for 529 (45.9%). Dose selection issues were the next most frequently reported, with 'dose too low' and 'dose too high' accounting for 13.2% and 10.8% respectively. The proportion of MSP MRPs in each subcategory was tested for evidence of statistically significant differences between study sites (to inform potential

generalisability of the MOAT); no statistically significant differences were found (given the Bonferroni corrected *p* value of 0.0022).

Table 35 – Classification of moderate or severe preventable medication related problems

Medication related problem (MRP) subcategory	**Hospital A (MSP MRPs = 757)** n (% of MSP MRPs)	**Hospital B (MSP MRPs = 396)** n (% of MSP MRPs)	**All patients (MSP MRPs = 1,153)** n (% of MSP MRPs)	***p* value (test for difference between study sites)**
1. Drug selection				
1.1 Inappropriate drug	41 (5.4)	22 (5.6)	63 (5.5)	0.921 (Chi-square)
1.2 No indication for drug / duplication	16 (2.1)	16 (4.0)	32 (2.8)	0.059 (Chi-square)
1.3 Interaction (drug-drug, or drugs and food / alcohol)	16 (2.1)	9 (2.3)	25 (2.2)	0.860 (Chi-square)
1.4 Indication not treated / missing therapy	351 (46.4)	178 (45.0)	529 (45.9)	0.646 (Chi-square)
1.5 More cost effective drug available	0	0	0	N/A
1.6 Synergistic / preventive drug required and not given	3 (0.4)	1 (0.3)	4 (0.4)	1.0 (Fisher's exact)
2. Drug form				
2.1 Inappropriate or suboptimal drug form	12 (1.6)	5 (1.3)	17 (1.5)	0.666 (Chi-square)
3. Dose selection				
3.1 Drug dose too low	99 (13.1)	53 (13.4)	152 (13.2)	0.884 (Chi-square)
3.2 Drug dose too high	79 (10.4)	45 (11.4)	124 (10.8)	0.629 (Chi-square)
3.3 Dosage regimen not frequent enough	1 (0.1)	1 (0.3)	2 (0.2)	1.0 (Fisher's exact)
3.4 Dosage regimen too frequent	4 (0.5)	2 (0.5)	6 (0.5)	1.0 (Fisher's exact)
3.5 Dose needs adjustment to organ function or change in disease state	20 (2.6)	5 (1.3)	25 (2.2)	0.127 (Chi-square)
3.6 Dosage instructions unclear, incomplete or not understood by patient / carer*	N/A*	N/A*	N/A*	N/A*
4. Treatment duration / withdrawal				
4.1 Duration of treatment too short	2 (0.3)	0	2 (0.2)	0.549 (Fisher's exact)
4.2 Duration of treatment too long	10 (1.3)	7 (1.8)	17 (1.5)	0.550 (Chi-square)
4.3 Inappropriate abrupt withdrawal†	2 (0.3)	0	2 (0.2)	0.549 (Fisher's exact)
5. Drug use process				
5.1 Inappropriate timing of administration / dosing by prescriber; administration error by nurse	8 (1.1)	3 (0.8)	11 (1.0)	0.620 (Chi-square)
5.2 Drug underused / under-administered	8 (1.1)	5 (1.3)	13 (1.1)	0.753 (Chi-square)

Continued from previous page…

Medication related problem (MRP) subcategory	**Hospital A (MSP MRPs = 757)** n (% of MSP MRPs)	**Hospital B (MSP MRPs = 396)** n (% of MSP MRPs)	**All patients (MSP MRPs = 1,153)** n (% of MSP MRPs)	***p* value (test for difference between study sites)**
5.3 Drug overused / over-administered	0	0	0	N/A
5.4 Drug not taken / administered at all	5 (0.7)	6 (1.5)	11 (1.0)	0.156 (Chi-square)
5.5 Wrong drug taken by patient	0	0	0	N/A
5.6 Drug abused	0	0	0	N/A
5.7 Patient or nurse uses drug incorrectly through lack of knowledge or barriers (e.g. swallowing, dexterity)	0	0	0	N/A
5.8 Adequate information not provided or not understood or misunderstood or not followed*	N/A*	N/A*	N/A*	N/A*
5.9 Drugs stored inappropriately / expired drug administered / preparation error	0	3 (0.8)	3 (0.3)	0.040 (Chi-square)
6. Logistics				
6.1 Prescribed drug not available	11 (1.5)	5 (1.3)	16 (1.4)	0.792 (Chi-square)
6.2 Drug order incorrect, incomplete, poorly legible / illegible / illegal / incorrect / allergy status incomplete	51 (6.7)	27 (6.8)	78 (6.8)	0.959 (Chi-square)
6.3 Error in drug selection	15 (2.0)	2 (0.5)	17 (1.5)	0.048 (Chi-square)
7. Monitoring				
7.1 Monitoring too frequent	0	0	0	N/A
7.2 No or too infrequent monitoring	3 (0.4)	0	3 (0.3)	0.555 (Fisher's exact)
7.3 Inappropriate test ordered	0	0	0	N/A
7.4 Patient unable to attend / pay for monitoring*	N/A*	N/A*	N/A*	N/A*
8. Unexpected reaction / adverse drug reaction (ADR) / no obvious cause				
8.1 An ADR occurred	0	1 (0.3)	1 (0.1)	0.343 (Fisher's exact)
8.2 No obvious cause of treatment failure	0	0	0	N/A

* Subcategory not used for MOAT study as relates to primary care (discussed in section 6.2.3)

† Subcategory not included in Basger's original classification system[114] (discussed in section 6.2.3)

MSP MRP = moderate or severe preventable medication related problem, N/A = not applicable

Bonferroni adjusted *p* value used to judge statistical significance 0.0022 (based on 23 statistical tests)

The overall results (combining MSP MRPs from both study sites) are comparable with the descriptive analysis of 'all MRPs' (Table 24, chapter 6), with 'indication not treated /

missing therapy' and issues related to dose selection being the most frequently identified for both sets of data. Slight differences may be explained by some MRPs being inherently less likely to result in adverse patient outcomes. For example, MRPs categorised as 'more cost effective drug available' are more likely to be associated with economic rather than patient safety benefits, therefore are unlikely to be rated as 'moderate or severe' in terms of potential adverse patient outcomes.

While no statistically significant differences were found between study sites in the classification of MSP MRPs, the analysis of 'all MRPs' found evidence for statistically significant differences for two subcategories ('indication not treated / missing therapy' and 'error in drug selection'), which appeared to relate to differences in the proportion of patients undergoing medicines reconciliation, and the use of different medication prescribing systems at the two sites. As above, this may suggest that the occurrence of MSP MRPs is less sensitive to differences in these hospital processes compared to 'all MRPs'.

8.3.3 Prevalence of outcome events

As shown in Table 36, 610 (40.6%) of 1,503 study admissions experienced the outcome event, namely at least one MSP MRP. Of the admissions that experienced an outcome event, the number of MSP MRPs per admission ranged from one to ten, with a median of one.

Table 36 – Prevalence of outcome event

Moderate or severe preventable medication related problem (MSP MRP) characteristics	**Hospital A (admissions = 1,006)** n (% of admissions)	**Hospital B (admissions = 497)** n (% of admissions)	**All patients (admissions = 1,503)** n (% of admissions)	***p* value (test for difference between study sites)**
Admissions with outcome event (at least one MSP MRP)	391 (38.9)	219 (44.1)	610 (40.6)	0.054 (Chi-square)
Number of MSP MRPs per patient (of the admissions with outcome event)	Median: 1 IQR: 1-2 Range: 1-10	Median: 1 IQR: 1-2 Range: 1-7	Median: 1 IQR: 1-2	0.287 (Mann-Whitney

Bonferroni adjusted *p* value used to judge statistical significance 0.025 (based on 2 statistical tests)

While there was evidence for a difference in the proportion of admissions who experienced at least one MRP between the two study sites (discussed in section 6.4.5.3), there was no statistically significant difference in the proportion of patients who experienced the outcome event (at least one MSP MRP) between study sites (p = 0.054), particularly given the Bonferroni adjusted p value of 0.025).

8.4 Discussion

Key findings

Of the 2,736 MRPs reviewed by the expert panel, 1,153 met the requirements for the outcome event, namely MSP MRPs. The majority of MSP MRPs were identified during (or before) the first review of the patient (73.9%). This compares to 15.3% that were identified during the remainder of the inpatient admission, and 10.6% identified during clinical screening of discharge prescriptions. The most frequently identified subcategories were 'indication not treated / missing therapy', 'dose too low' and 'dose too high', accounting for almost 70%, with no statistically significant differences in the proportion of MRP subcategories between study sites. Six hundred and ten (40.6%) of the 1,503 study admissions experienced an outcome event, namely at least one MSP MRP.

Comparison with previous literature

No estimation of the prevalence of the MSP MRPs in the United Kingdom exists, but Blix *et al*[78] (Norway, 2004) assessed the 'clinical significance' of MRPs for a subset of patients included in a study to describe the frequency and types of MRPs in hospitalised patients. Blix *et al* reviewed the MRPs for every sixth patient in their dataset, and found that 5.9% of MRPs were 'extremely important' (with the potential to cause death, severe or irreversible detrimental effects), 43.7% were 'major' (requiring intervention to prevent major or reversible detrimental effects, or lack of therapy), 40% were 'moderate' (where intervention resulted in moderate benefit), and 10.4% were of 'minor' clinical significance. Although this grading system is not directly comparable with the present study, if we assume a moderate or severe MRP (graded using the visual analogue scale[163] used for the MOAT study), equates to an 'extremely important' or 'major' MRP (using Blix's definitions), the proportions are broadly comparable; Blix *et al* reported that 49.6% of MRPs were either extremely important or major, compared to 44.2% of MOAT MRPs being found to be moderate or severe (i.e. 1,153 of the 2,609 preventable MRPs).

Strengths and limitations

Strengths of the MRP severity rating include use of a validated method (used previously for this purpose), and the selection of clinically experienced expert panel members from a range of professions, increasing the credibility of the assessment and facilitating ownership of the results by the professions involved.

A potential limitation was the inclusion of myself on the expert panel, as I could not be fully blinded to the candidate predictors, and had a potentially vested interest in the grading of MSP MRPs. This was necessary due to financial constraints, but had the potential to bias results in two ways. First, access to candidate predictor data meant imperfect blinding, despite the use of anonymisation. Second, there was potential for confirmation bias due to a subconscious desire to influence the number of MSP MRPs[154]. This was addressed in the following ways:

- separation of MRP data entry / severity rating and candidate predictor data collection by as much time as possible (to reduce the likelihood that I may recognise study codes, or recall related data). In addition I avoided accessing candidate predictor data during MRP severity rating;
- use of a review / discussion of 35 MRPs by panel members (prior to carrying out the MRP assessment). This permitted identification of potential discrepancies in panel member's approaches to grading, and led to the development of simple conventions to standardise severity assessments between myself and other panel members. The 'sample grading' provided an initial benchmark, then to minimise 'judgement drift' I referred to it throughout the assessment period, and kept an on-going 'case law document'[143].

Another potential limitation was the occurrence of a small amount of missing data, with the 'stage of patient stay' missing for three (0.3%) of 1,153 MSP MRPs. As this data was collected for descriptive purposes only, it did not impact on development of the MOAT.

As in previous chapters, the limitations associated with the use of the Bonferroni correction also need to be considered[146]. The Bonferroni correction was used as the primary objective was to test universal null hypotheses, and as previously, false positive errors were a greater issue than false negative errors (as the purpose was to establish an answer to these hypotheses, rather than generate hypotheses for further investigation).

Implications

Given the above limitations, it is possible that subsequent studies may be able to improve on the research presented in this chapter by excluding members of the research team from the expert panel, and collecting MRP data under strict trial conditions (to avoid ambiguities or missing data).

The findings of this chapter provide additional support for the potential generalisability of the MOAT, in that no statistically significant differences were observed in the prevalence of the outcome event (i.e. the proportion of study admissions that experienced at least one MSP MRP), or 'outcome components' (i.e. the MSP MRP subcategories) between study sites[60]. While this only represents two hospitals, both in the South East of England, it does provide some evidence for consistency. Indeed, the comparison between analyses for 'MSP MRPs' and 'all MRPs' found that differences between study sites were less pronounced for 'MSP MRPs', suggesting that differences in hospital processes, that appear to impact on the prevalence of 'all MRPs', have less influence on the occurrence of 'MSP MRPs'.

Analysis of the stage in admission when MSP MRPs were identified suggests that the MOAT should be used to target patients at the point of admission to hospital. Results found that the majority of MSP MRPs (73.9%) were identified during (or before) the first review of the patient, with 52.4% related to medicines reconciliation discrepancies (often undertaken during the first review). Future research may be warranted to investigate whether this is influenced by working practices at the study sites (for example, it may reflect a focus on new admissions), but based on the information currently available, these results suggest that patients are at highest risk of experiencing MSP MRPs in the early stages of their hospital stay. Regarding subsequent stages in hospital stay, 15.3% of MSP MRPs occurred during the 'remainder of inpatient stay' and 10.6% during clinical screening of discharge prescriptions, suggesting that risk may diminish, but that patients continue to require pharmacist review. Risk is unlikely to be static throughout admission, in practice risk may reduce because MRPs have already been identified and resolved, fewer prescribing alterations occur, or due to changes in modifiable predictors, for example the discontinuation of high-risk or parenteral medicines, or improvements in laboratory results. Similarly, risk may increase as a patient's treatment changes or their condition deteriorates. A potential limitation of the MOAT is that it has been designed to identify patients at risk of experiencing an outcome event *at any stage* in the hospital stay, that is, the predictions are based on 'startpoint data' (i.e. on admission to a hospital medical ward), meaning it is not possible to predict if or when a patient's risk status will change. A future development could be to investigate whether predictors vary depending on the stage of patient stay, potentially leading to the development of separate risk prediction tools that are specific to the different stages. This would require a larger sample than the present study as sufficient MSP MRPs for each stage of patient stay would be needed to permit robust statistical modelling. Repeated measurement for modifiable

candidate predictors would also be needed (to provide startpoint data for each stage of patient stay).

The results may also be of use to future MOAT users when assessing the applicability of the MOAT, as they provides an indication of working practices at the study sites; for example, the high proportion of MSP MRPs identified during routine ward visits (93.6%) and resolved by pharmacy staff (98.2%).

8.5 Conclusion

A validated method was used to grade the severity of study MRPs[163], and identified 1,153 that met the definition for the outcome event (MSP MRPs). This equates to 44.2% of the total MRPs, which is consistent with previous literature[78]. In terms of the prevalence of MSP MRPs, 610 (40.6%) of 1,503 study admissions experienced at least one MSP MRP, but this may underestimate the true prevalence (given that study pharmacists identified approximately 85% of simulated MRPs as part of an identification assessment exercise). Further analysis found there were no statistically significant differences in the proportion of each subcategory of MSP MRPs between study sites. MSP MRPs were most frequency identified during or before the first review by pharmacists, suggesting the MOAT should be used to target high-risk patients as early as possible following admission to hospital.

The next chapter will describe the development of the MOAT. This will involve the use of multivariable logistic regression to determine the relationship between candidate predictors and patients with an outcome event.

Chapter 9: Modelling

9.1 Introduction

As discussed in chapters 2 and 6, recommendations to improve the quality and impact of prognosis research were published by the PROGnosis RESearch Strategy (PROGRESS) group in 2013[53]. In relation to the quality of studies, they recommended the need for 'integrated standards of design, analysis and reporting', with specific advice on the need to address deficiencies in statistical modelling, including: 'multiple sources of significance chasing bias, lack of appreciation of type II errors associated with small sample sizes, and the arbitrary dichotomisation or categorisation of continuous variables'. The 'CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies' (CHARMS)[60], and 'Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis' (TRIPOD) statement[57] followed; providing further guidance on the qualities of good studies (and inappropriate approaches that should be avoided), potential sources of bias, and the level of reporting required to permit quality to be adequately assessed.

The aim of the work presented in this chapter was therefore to develop the Medicines Optimisation Assessment Tool (MOAT^{TM}) using approaches informed by the PROGRESS[53 55], TRIPOD[94], and CHARMS[60] recommendations. The objectives were to:

- select statistical methods that minimise the risk of bias;
- develop a regression model that predicts the outcome event well with the minimum number of variables;
- develop a decision aid (the MOAT) to permit pharmacists to target patients at highest risk of experiencing the outcome event;
- report all stages of the statistical analysis clearly and fully (to permit assessment of the quality of the study to be assessed).

9.2 Methods

The methods are presented according to the TRIPOD reporting guidelines[57]. This section includes:

- handling of continuous predictors;
- pooling of data;
- sample size calculation;
- exploratory data analysis;
- imputation of missing data;
- model development;
- assessment of model performance;
- internal validation / adjustment for model 'optimism';
- development of the MOAT (decision aid based on the regression model).

Each of these will be described in turn; results are described in section 9.3.

9.2.1 Handling of continuous predictors

All continuous predictors were analysed on their continuous scale, that is dichotomisation or categorisation were not used (as this can lead to optimistic model performance[58 60 62 94]). As scaling of continuous variables is important for interpretability and comparability of effects[70], predictors with a wide range in units (age, socioeconomic status, estimated glomerular filtration rate, and platelet count) were analysed as deciles (i.e. after dividing the actual value by ten). This approach is recommended by Steyerberg, on the basis that changes in predictive effect per one unit increase would otherwise be small, making comparison with other predictors more difficult[70].

9.2.2 Pooling of data

Given that no statistically significant differences were found in the prevalence of outcome events between study sites (reported in chapter 8), data from the two study sites were combined for the exploratory data analysis and model development.

9.2.3 Sample size calculation

Sample size is often calculated based on power calculations, but this is not straightforward for prognostic modelling studies as there is often not a clear 'measure of effect' to power the research. An alternative method is to calculate the sample size based on the desired precision of a sample estimate[94]. An alternative approach that is commonly used is the 'rule of thumb' of at least 'ten events per variable' (EPV)[92]. This method requires the sample size to be based on the prevalence of the outcome event and the number of candidate predictors that will be used in model development[60 92-94]. Although there is debate over the optimal number of EPV, with recognition that 'the rule of ten or more EPV is not a well-defined bright line'[95], there is agreement that models developed with less than ten EPV need to be interpreted with caution[94 95]. The reason for the potential problem with using less than ten EPV relates to the reliability of the model when used in a new group of patients. If a model is too closely adapted to the developmental data it can reflect associations between the candidate predictors and outcome event that are due to chance rather than true associations, known as 'overfitting' or 'optimism'[60 62 94].

For the present study the sample size was dictated by practical considerations (funding, time available, and accessibility of data at the study sites), resulting in the capacity to include 1,500 admissions. I therefore used both the precision and EPV methods to consider the adequacy of this sample size, based on an estimation of the outcome event prevalence in the study population.

The outcome event of interest for this study was the occurrence of at least one moderate or severe preventable medication related problem (MSP MRP) in hospitalised UK patients. As discussed in chapter 8, no estimate for the prevalence of this outcome event currently exists, but Blix *et al*[78] (Norway 2004) reported that 81% of 827 hospitalised patients experienced an MRP, with approximately half of all MRPs classified as 'extremely important' or 'major' in terms of clinical significance (preventability not reported). To establish the prevalence of the outcome event in the study population I carried out pilot work involving 200 patients, and found that 39% (95% confidence interval 32-45%) experienced at least one MSP MRP (assessed using the same methods used for the main study). Although this was consistent with Blix's work, the estimate was based on a small sample of 200 patients. In addition the MRPs were severity-rated by three members of the expert panel, rather than the four used for the main study. I therefore chose to use the lower confidence interval limit as an estimate of event prevalence, i.e. 32%. Given an anticipated event prevalence of 32%,

and a sample size of 1,500, it was anticipated there would be 480 admissions with at least one MSP MRP.

To consider the adequacy of the sample size using the precision method I first established acceptable target sensitivity for the MOAT by including a question in the survey of healthcare professionals and patient / public representatives (discussed in chapter 5). A target sensitivity of 90% was proposed, and survey respondents asked if this was acceptable. This sensitivity was selected based on previous research to develop a 'clinical decision rule' to identify emergency department patients at risk of adverse drug events[169]. Hohl *et al* used a target sensitivity of 90% as this was deemed acceptable by emergency physicians, and considered feasible for implementation in terms of workload for pharmacists[169]. A total of 237 responses were received for this question: 189 (80%) answered that 90% was an acceptable target; 21 (9%) answered no, and 27 (11%) were 'unsure'. This was discussed with my academic supervisors, and it was concluded that 90% was an acceptable target for sensitivity. Given the anticipated number of outcome events and a target sensitivity of 90%, this would permit the precision of the sensitivity to be estimated with 95% confidence intervals of $\pm$2.7%, which we considered to be an acceptable level of precision in terms of clinical usefulness of the MOAT.

For the EPV method the aim would be to have at least ten events for every variable used in model development. Given the estimate of 480 outcome events, this would permit the inclusion of 48 'variables' in model development. The number of variables includes all proposed candidate predictors, interactions examined (that is where a candidate predictor has a different association with the response variable depending on the value of a third variable), transformations for continuous predictors (which permits modelling of non-linear predictors), and indicator variables for categorical predictors.

As discussed in chapters 5 and 6, the candidate predictors were pre-selected to reduce the risk of selection bias[62]. Table 37 shows the total number of variables that were proposed for each pre-selected candidate predictor (including the indicator variables, i.e. the artificial variables used to represent distinct groups within a categorical variable), giving a total of 37 variables. This resulted in 13 EPV assuming no interactions or transformations were required (it was not possible to predict if transformations would be required before examining the linearity of continuous predictors, and no interactions were hypothesised *a priori*).

Table 37 – Number of variables for development of the Medicines Optimisation Assessment Tool based on pre-selected candidate predictors

Variable	Type of measurement	Number of variables*
Age	Continuous numeric	1
Socioeconomic status	Continuous numeric	1
Previous allergy / adverse drug reaction	Binary (YES/NO)	1
Body mass index	Continuous numeric	1
Number of hospital admissions in previous 6 months	Continuous numeric	1
Primary diagnosis	Nominal categorical (7 categories)	6
Number of comorbidities	Continuous numeric	1
History of dementia	Binary (YES/NO)	1
Number of medicines prescribed	Continuous numeric	1
Parenteral administration route	Binary (YES/NO)	1
Use of 'high-risk medicines'	Binary (YES/NO) for each of 15 groups	15
Renal function (glomerular filtration rate calculated using the modified Modification of Diet in Renal Disease equation)	Continuous numeric	1
Liver disease	Binary (YES/NO)	1
Serum albumin	Continuous numeric	1
Serum potassium	Continuous numeric	1
Serum sodium	Continuous numeric	1
White cell count	Continuous numeric	1
Platelet count	Continuous numeric	1
Total number of variables*		37

* Number of variables in relation to calculating the 'events per variable'

Review of adequacy of sample size following data collection

Following data collection it was possible to review the adequacy of the sample size based on the actual prevalence of the outcome event in the study sample (40.6%, i.e. 610 of 1,503 admissions experienced at least one MSP MRP).

Using the precision method, the higher number of events (compared to the provisional estimate) equates to an increased precision in the estimation of the MOAT's sensitivity (95% confidence intervals $\pm$2.4%).

Two changes impact on the EPV calculation:

- increased number of admissions experiencing the outcome event compared to provisional estimate (610 versus 480);

- reduced number of variables from 37 to 33 (due to the need to collapse four of the proposed high-risk medicine categories, discussed further in section 9.3.1.1).

As non-linear transformations were not required (discussed in section 9.3.1.3), and no interactions examined, this resulted in an increase in the EPV to 18. While the original estimate was adequate (i.e. above ten EPV), it is acknowledged that more events and higher EPV are almost always prefereable[95] due to the reduced risk of model overfitting[60 62 94].

9.2.4 Exploratory data analysis

Descriptive statistics for the candidate predictors were presented in chapter 6, including percentages for categorical predictors, and measures of central tendency, variability, and ranges for continuous predictors. Further exploratory data analyses were carried our prior to undertaking model development to maximise insight into the dataset, specifically to:

- identify categorical predictors with narrow distributions (resulting in the need to collapse groups);
- identify values of continuous predictors outside their typical ranges (i.e. outliers);
- assess linearity between continuous predictors and the outcome event (to identify the need to include non-linear transformations during multiple imputation and model building);
- test for multicollinearity between candidate predictors;
- establish univariable associations between predictors and the outcome event.

Each of these will be described in turn.

Stata version 14.2 was used for all statistical analyses using the observed (i.e. non-imputed) data.

9.2.4.1 Review of distribution of categorical predictors

It was necessary to review the categorical candidate predictors prior to data analysis to identify and eliminate sparse categories[94]. This prevents low predictor prevalence biasing results, or causing non-convergence of regression models[63 95]. Although there is no clear definition of 'sparse', 5% has been used as the minimum predictor prevalence by other researchers developing prognostic models for adverse medication-related outcomes[42 44 45].

Ideally the decision to collapse categories should be made while blinded to the relationship between predictor and outcome events in the sample population, as this limits overfitting[70 94]; the prevalence of all categorical predictors was therefore reviewed prior to univariable or multivariable regression modelling. Predictor categories containing less than 5% of the study population were identified, then consideration given to alternative categorisation, collapsing categories only where clinically sensible.

9.2.4.2 Identification and review of outliers (continuous predictors)

Outliers, defined by Tukey as values more than (or equal to) one and a half interquartile ranges (IQRs) above the third quartile or below the first quartile[170], have the potential to substantially distort statistical estimates and inflate error rates[171]. Box-plots were therefore used to identify outliers, and where identified, values were reviewed to establish plausibility. Implausible values were considered as errors and hence set to missing[70].

As advised by Steyerberg[70], 'truncation' was used to reduce the influence of outliers on the regression coefficients (known as leverage). This involved assessing the impact of shifting very high and very low values to truncation points, defined by examining the distribution of data, and predictor-outcome relationship. More specifically:

- truncation points were set at one and a half times the IQRs above or below the third or first quartile respectively, with truncation points for laboratory results checked against standard reference ranges (to ensure clinical sensibility);
- outliers were identified using visual inspection of box-plots, and Stata's 'extremes' module[172] (which provides distance of outliers from the nearer quartile);
- univariable logistic regression models were specified for each predictor using non-truncated and truncated data. Where predictors had outlying values that were both above the upper truncation point, and below the lower truncation points, the impact of each group of outlying values were considered separately;
- 'regression coefficients' (i.e. the relationship between independent and dependent variable, more specifically the change in predicted log odds[iii] of the regression outcome that would be predicted by a one unit increase in the candidate predictor) were then compared to identify 'substantial differences' between the non-truncated and truncated models. As there is no accepted definition of what constitutes a 'substantial difference', I chose to interpret a change in regression coefficients of 10% or more as being substantial. This was to permit a consistent approach, and reduced the risk of significance chasing bias[53] (i.e. selection based on which version of the data produced the most statistically significant model). Where there was no substantial difference in the regression coefficients between the model fitted with non-truncated and truncated data, non-truncated data were used for subsequent analyses and MOAT development (on the basis that the outlying cases

[iii] Logistic regression uses a combination of odds ratios and natural log transformations, therefore the probability of the outcome event is modelled on a log-odd scale. This permits predictions to be unbounded (i.e. from plus infinity to minus infinity), preventing invalid values (such as probabilities above one or below zero).

were representative of the population of interest[173]). Where a substantial difference was observed, truncated data were used.

The model chi-square for non-truncated and truncated models was also compared to indicate the impact of truncation on model fit. The model chi-square is twice the difference in the log likelihood of a model with and without the predictor(s)[70], and is an omnibus test of statistical significance of the model[174]. It has no meaning in itself, but provides a comparison of nested models[175]. As the truncated and non-truncated models were fitted using slightly different data, it is not possible to directly compare their log likelihoods using standard statistical tests such as the likelihood ratio test (which compares the log likelihoods of two models under the null hypothesis of no association[70]); model chi-square results were therefore recorded to provide an *indication* of the change in the predictive information provided by the predictor before and after truncation, with a decrease indicating a loss of predictive information[70].

Truncation was used rather than data transformation (for example using logarithms or quadratics), due to the potential impact of transformations on the interpretability of the MOAT (as regression coefficients would no longer relate to meaningful measurement scales)[171].

9.2.4.3 Linearity

Standard logistic regression models assume linearity between continuous predictors (modelled on their continuous scales) and the log-odds of the outcome variable, that is to say the effect is the same at each part of the predictor range[70]. Where this is not the case, it is necessary to account for the non-linearity during model building. While checks for non-linearity are often performed during model development, it was necessary to identify potential non-linearity prior to imputing missing data, to permit the inclusion of non-linear transformations in the multiple imputation model; failure to do so can cause misspecification of the imputation model, leading to biased results[176 177].

Non-linearity can be checked using various methods, such as applying simple transformations (for example, logarithms or square roots), the inclusion of quadratic or cubic polynomials as an extension to a model (for example $x,\ x + x^2,\ x + x^2 + x^3$), or the use of fractional polynomials (which extends ordinary polynomials by including non-positive and fractional powers)[70]. Improvements in model fit are then investigated (i.e. the extent to which fitted values of the outcome variable compare with observed values[178]) following inclusion of non-linear transformations in the model.

Traditionally quadratic or cubic polynomials have been used to check for non-linear relationships, but they are limited in the shapes they can take. Fractional polynomials (FPs) allow for smoother and more flexible transformations, therefore are now widely recommended[70 94]. It is also possible to use FPs to assess the linearity of all continuous predictors simultaneously, using 'multivariable fractional polynomial (MFP) modelling'[179 180]. Given these advantages, I decided to use MFP modelling to check for non-linearity between the candidate predictors and outcome.

In MFP modelling, null models (that exclude each predictor in turn), are compared to models including the predictor as a linear variable, and models containing the predictor with one or two FPs (selected by searching all power combinations from the set -2, -1, -0.5, 0, 0.5, 1, 2, 3, where inverse is x^{-1}, square root is $x^{0.5}$, log is x^{0}, linear is x^{1}, squared is x^{2}, and so on). Where there is evidence for non-linearity of the continuous predictor (from comparisons of model fit), either one or two FP transformations are automatically selected in the final MFP model, depending on which fits the data better.

The Stata code used to run the MFP model was:

```
. xi: mfp, alpha(.05) select(1): logit study_outcome age_dec_F ses_dec_F bmi_trun_F num_hosp_adm_trun_F num_comorb_F num_meds
> _F egfr_trun_F albumin_trun_F potassium_trun_F sodium_F wcc_trun_F platelets_trun_F allergy i.prim_diag dementia iv_use ant
> icoags heparin diabetes opiates gent_vanc antimicrob epilepsy antipsychot antiarryth antidepres other_high_risk liver_dx
```

where:

- 'xi:' expands terms containing nominal categorical variables (i.e. primary diagnosis);
- 'mfp' is the command to perform MFP modelling;
- 'alpha' sets the significance levels for testing between different fractional polynomials models (0.05 is the default nominal *p* value);
- 'select' sets the nominal significance level for variable selection by backward elimination. This was set at 1 to force all variables to remain in the model (i.e. to produce a full model);
- 'logit' is the command to run a logistic regression;
- the logit command is followed by the outcome variable, then a list of all candidate predictors (e.g. 'age_dec_F' was the variable name used for 'age in deciles').

I also ran a MFP model using the backward elimination option to determine if linearity was altered following removal of 'non-significant' variables. This involved changing the 'select' option to 0.157 (i.e. the nominal significance level used for development of the MOAT, discussed further in section 9.2.6.2).

9.2.4.4 Multicollinearity

Multicollinearity occurs when two or more candidate predictors are highly correlated[181]. This can cause problems when fitting and interpreting regression models, as very strong correlations can prevent accurate estimation of the regression coefficients of affected predictors[70]. As a result, Steyerberg advises that where predictors are strongly correlated, it may be appropriate to combine them into a single combined variable[70]. To establish whether this would be necessary for the present study I assessed multicollinearity by calculating variance inflation factors (VIFs), which measure 'the degree to which collinearity among the predictors degrades the precision of estimate coefficients'[70].

Where predictors are completely uncorrelated with each other, the VIF is equal to one, whereas if predictors are very closely related to other variable(s) the VIF gets very large[174]. A rule of thumb is that a VIF greater than ten indicates multicollinearity[70].

The Stata programme 'collin' was used to calculate the VIF for each predictor, using the following code:

```
. xi: collin age_dec_F ses_dec_F bmi_trun_F num_hosp_adm_trun_F num_comorb_F num_meds_F egfr_trun_F albumin_trun_F potassium_
> trun_F sodium_F wcc_trun_F platelets_trun_F allergy i.prim_diag dementia iv_use anticoags heparin diabetes opiates gent_van
> c antimicrob epilepsy antipsychot antiarryth other_high_risk liver_dx
```

where:

- 'xi:' expands terms containing nominal categorical variables (i.e. primary diagnosis);
- 'collin' is the command to perform the multicollinearity diagnostics;
- the collin command is followed by a list of all candidate predictors (e.g. 'age_dec_F', the variable name used for 'age in deciles').

9.2.4.5 Univariable analyses

Univariable analyses were performed to provide data on the unadjusted association between candidate predictors and the outcome event. Cross-tabulation was used to compare admissions with and without the outcome event. Proportions were recorded for categorical predictors, and means recorded as a measure of central tendency for continuous predictors (as examination of the truncated data suggested compatibility with normal distributions). Odds ratios were also obtained using univariable logistic regression.

9.2.5 Missing data

The comparison of study admissions with and without missing values (chapter 6) suggested data were missing at random (MAR). Due to the risk of selection bias associated with the use of complete-case analysis when data are MAR[94 148], in addition to a loss of statistical power / precision associated with a reduced sample size[62], it was decided to deal with missing data using multiple imputation. The purpose of multiple imputation is to permit valid statistical inference, rather than to recreate individual missing values[176]. This is achieved by creating multiple datasets, with missing observations substituted with plausible values based on other observed participant data[60]. Analysis is then performed on each completed dataset and the results combined, so allowing for the uncertainty around the true values[148]. This differs from single imputation, where missing values are treated as known in subsequent analyses; single imputation therefore underestimates the variance of the estimates (by not taking sampling variability into account), and overestimates precision, resulting in confidence intervals and significance tests that are too optimistic[176].

In a review of the use of multiple imputation in epidemiology and clinical research, Sterne *et al*[148] advised that multiple imputation 'needs to applied carefully to avoid misleading conclusions'. Sterne proposes reporting guidelines for missing data (which extend the STROBE guidelines for observational studies[147], and TRIPOD statement for prognostic model studies[57]). These guidelines are summarised below:

1. report number of missing values, with reasons for missing values in terms of other variables;
2. clarify whether there are important differences between individuals with and without missing data (e.g. comparison table of key variables);
3. describe the type of analysis used to account for missing data, the missingness assumption used, and for multiple imputation, whether the MAR assumption is plausible;
4. for analyses based on multiple imputation provide:
    - software used and key settings for imputation modelling;
    - number of imputed datasets created (with at least 20 preferable);
    - which variables were included in imputation procedure;
    - how non-normally distributed and categorical variables were dealt with;
    - whether interactions were included in final analyses and imputation models;
5. comparison of observed and imputed values (where large fraction of data imputed);

6. sensitivity analysis comparing results from complete-case analysis with those from multiple imputation, with investigation of 'the robustness of key inferences to departure from the MAR assumption'.

Items one and two are reported in chapter 6. Regarding item 3, the analysis of the likely missingness mechanism, and plausibility of the MAR assumption are also reported to chapter 6. As a result of these assessments, multiple imputation was chosen to deal with missing data, but prior to undertaking the imputation modelling I explored the data to identify any missing data that could be predicted using less formal methods[182] (results presented in section 9.3.2.1). Items four to six are addressed below, together with the 'imputation diagnostics', which were performed to assess how well the imputation performed.

9.2.5.1 Software and imputation method

Imputation modelling was conducted using Stata version 14.2. 'Multiple imputation by chained equations' (MICE) was used to create 30 imputed datasets. A random seed (53421) was used to permit results to be reproduced.

9.2.5.2 Selection of variables for imputation modelling

Sterne *et al*[148] advise that the MAR assumption is only valid if variables predictive of the missing values are included in the imputation model, and that bias is only avoided if sufficient variables (that are predictive of the missing values) are included. Sterne recommends inclusion of all candidate predictors, plus 'auxiliary variables', which are variables predictive of the missing values, and those influencing the process causing the missing values. The following variables were therefore included in the imputation model as auxiliary variables:

- weight and height (as predictive of BMI for those patients with only one or other measurement available);
- gender (as being male appears to be positively associated with missing data, and gender is likely to be associated with BMI);
- length of hospital stay (as shorter hospital stay was associated with missing data).

Pairwise correlation supported use of these variables, with positive correlations between weight and height (r = 0.46), weight and gender (r = 0.27), weight and BMI (r = 0.87), height and gender (r = 0.65), and length of hospital stay and serum albumin (r = 0.30).

As recommended by Sterne *et al*[148], the outcome event was also used to impute the missing predictor values. This is advised as the outcome event carries information about the missing predictor values, and failure to include it may falsely weaken associations between predictors and outcomes.

Truncated candidate predictor data were used for the imputation model where selected for the MOAT development model (as discussed in section 9.2.4.2). This was to preserve data characteristics that would later be explored at the analysis stage[176].

9.2.5.3 Analysis of data distribution

Sterne *et al*[148] advise that the inclusion of non-normally distributed variables can lead to bias and/or implausible results, therefore the distributions of continuous variables were established by plotting the data and reviewing the skewness statistic, which is a measure of the degree and direction of asymmetry, with symmetric (normal) distribution having a skewness of zero. A skewness statistic of between minus one and plus one was interpreted as representing normal distribution[183]. Normality tests (such as the Shapiro-Wilk test) were not used as they can be overly conservative for large samples (i.e. greater than 300)[184]. Transformations were then applied as appropriate, with the choice dependent on the direction and extent of skew[181].

9.2.5.4 Inclusion of variable transformations

It is recommended that variable transformations intended for use in the final analyses, such as logs, FPs and interaction terms, are included in the imputation model to prevent misspecification of the model[176 177].

No interaction terms were included in the imputation model as none were hypothesised *a priori* for the final analyses. Regarding non-linear transformations, the exploratory data analysis found no evidence for non-linearity between the continuous predictors and outcome variable (section 9.3.1.3); it was therefore not necessary to include non-linear transformations in the imputation model or final analyses.

9.2.5.5 Comparison of observed and imputed values

The MICE imputation model was specified using linear regression as the conditional distribution for variables with missing values (as all variables with missing data were continuous)[176 177]:

```
. mi impute chained (regress) ses_dec_F bmi_trun_F egfr_trun_dec_F albumin_trun_F potassium_trun_F sodium_F wcc_trun_F platel
> ets_trun_dec_F weight height = age_dec_F num_hosp_adm_trun_F num_comorb_F num_meds_F i.prim_diag allergy dementia liver_dx
> iv_use anticoags heparin diabetes opiates gent_vanc antimicrob epilepsy antipsychot antiarryth antidepres other_high_risk s
> tudy_outcome ln_length_stay male , add (30) rseed (53421) savetrace(trace1,replace)
```

where:

- 'mi impute chained' is the Stata command to run a MICE imputation;
- 'regress' specifies that MICE distribution for imputation of the missing variables (i.e. linear regression);
- variables on the left of the equal sign have missing information, those on the right have no missing values (e.g. 'ses_dec_F' was the variable name used for 'socioeconomic status in deciles').;
- 'add' specifies the number of imputations to be performed (i.e. 30 imputations);
- 'rseed' is the random seed used to permit reproducibility of the results;
- 'savetrace(trace1,replace)' specifies to Stata to save a file of the predicted values from each iteration of the imputation for the purposes of imputation diagnostics (section 9.2.5.6).

All imputed values were then reviewed to identify implausible results (e.g. negative or clinically implausible values). Where the imputed values were implausible, truncated distributions were specified (i.e. upper and/or lower limits) using Stata's 'truncreg' imputation method[176].

9.2.5.6 Imputation diagnostics

It is possible to assess how well an imputation has performed by fitting a standard logistic regression model using the imputed data[176 177], then checking the:

- Relative Increase in Variance (RVI) – the proportional increase in total sampling variance that is due to missing information;
- Fraction of Missing Information (FMI) – the proportion of the total sampling variance that is due to missing data.

The RVI is interpreted by how close it is to zero (signifying minimal effect of missing data on the variance of the estimate). FMI is used to assess whether sufficient imputations have been used, with a commonly used rule of thumb that the number of imputations should be at least equal to the highest FMI percentage.

I also checked convergence of the MICE algorithm by visually examining 'trace' plots for BMI and serum potassium, chosen as these variables had the highest proportion of missing values (22.7% and 2.0% respectively)[177]. The option 'savetrace' was used when imputing the missing values (as shown in section 9.2.5.5). This specifies Stata to save the means and standard deviations of imputed values from each iteration to a Stata dataset[177]. The predicted mean and standard deviation for BMI and serum potassium were then plotted, with the ten imputation chains (the default number used by MICE) graphed simultaneously to establish whether anything unexpected occurred in a single chain; absence of any sort of trend in the summaries of the imputed values indicates good convergence[177].

9.2.5.7 Multiple imputation sensitivity analysis

As discussed in chapter 6, multiple imputation is not appropriate where data are missing not at random (MNAR), as MNAR data are related to unobserved data, and therefore cannot be plausibly estimated from the observed study variables. Sensitivity analysis was therefore carried out to compare complete-case and multiple imputation analyses.

Multivariable logistic regression models (including all candidate predictors) were specified for the complete-case and imputed datasets; regression coefficients, standard errors, and *p* values were then compared. Given the large number of admissions with missing BMI data, I also re-ran the analyses, omitting BMI; this increased the number of observations for the complete-case analysis.

9.2.6 Model development

Model development is presented in three sections: the choice of model type, selection of predictors during modelling, and model diagnostics. Each of these is described in turn.

9.2.6.1 Type of model

Multivariable binary logistic regression was selected for model development. This was chosen as the outcome event is binary, and all participants were followed up to the end of the study period. A random effects model was used to account for possible correlation between observations, that is, between patients admitted more than once during the study period. Failure to take account of this lack of independence may have resulted in standard errors and *p* values being too small, hence confidence intervals too narrow, resulting in a belief that evidence was stronger than it actually is[185].

The full model was specified using the Stata code:

```
. mi estimate: xtlogit study_outcome age_dec_F ses_dec_F bmi_trun_F num_hosp_adm_trun_F num_comorb_F num_meds_F egfr_trun_dec
> _F albumin_trun_F potassium_trun_F sodium_F wcc_trun_F platelets_trun_dec_F allergy i.prim_diag dementia iv_use anticoags h
> eparin diabetes opiates gent_vanc antimicrob epilepsy antipsychot antiarryth antidepres other_high_risk liver_dx, re i(id)
```

where:

- ‘mi estimate’ runs the model estimation command on the multiply imputed data, adjusting coefficients and standard errors for the variability between imputations according to the combination rules by Rubin[176];
- ‘xtlogit’ is the command to run a ‘random effects’, ‘fixed-effect’, or ‘population-averaged’ logistic regression;
- the xtlogit command is followed by the outcome variable, then a list of all candidate predictors (e.g. ‘age_dec_F’ was the variable name used for ‘age in deciles’);
- ‘re’ is the option used to specify a ‘random effects’ model;
- ‘i(id)’ indicates the data structure (i.e. identifies the duplicate admissions).

9.2.6.2 Predictor selection during modelling

The aim was to produce a parsimonious model to increase clinical applicability while retaining reasonable predictive performance; it was therefore necessary to reduce the number of candidate predictors included in the final prediction model. Univariable associations between predictors and the outcome event were not used to preselect

variables; this is not recommended on the basis that important predictors may be excluded due to their predictive effect being masked by other predictors[94].

Backwards elimination was used to reduce the set of candidate predictors during modelling, with the aim of including only the most significant in the model. Backward elimination starts with a full model (i.e. one that contains all predictors), the 'least significant' predictor (based on a predetermined 'stopping rule'), is then removed, and the model re-fitted. This is continued until all predictors in the model are 'significant'[70]. I used a significance level of *p* greater than 0.157 to exclude predictors, chosen as it is comparable with the more complex Akaike Information Criterion (AIC)[70], which compares models based on their fit to the data while penalising for the complexity of the model. Use of the AIC, or even higher *p* values (for example a *p* value of greater than 0.5), is considered to be a suitable for relatively small dataset (hence relatively larger p-values for the predictors) as this is less likely to result in underfitting than alternative methods[70]. Automated variable selection was not used because: (1) manual selection permitted clinical judgement to be incorporated, for example to decide which predictor to exclude in cases where two predictors had similar significance levels; and (2) automated techniques are generally considered to have a high probability of generating spurious findings[63].

As 'primary diagnosis' was a nominal categorical variable it was necessary to select a 'base category'. This does not fundamentally alter the estimation, but provides a base level, with the reported coefficients for the remaining categories measured as the difference from this base category[186]. I chose to use the category 'other' (which consisted mainly of symptoms or findings not specific to a diagnosis) as this would be difficult to use as a predictive category in clinical practice. It also contained a reasonable number of admissions, that is 244 (16.2%) of 1,503 admissions, so increasing the reliability of parameter estimates[181]. The remaining primary diagnosis categories were then considered individually during the backwards elimination, and removed as appropriate.

Once all predictors in the model were significant (i.e. *p* smaller than or equal to 0.157), I reviewed the selected predictors to consider the sensibility of including each of them in the MOAT. This involved consideration of whether the predictor's contribution to the predictive ability of the model justified the workload / inconvenience of obtaining the data in clinical practice. The concordance index (c-index) was used to estimate predictive ability (discussed further in section 9.2.7).

Results are presented for the full regression model (i.e. containing all candidate predictors), and the model produced following backward elimination; the latter will be referred to as the 'backward selection model' (BS model) to distinguish it from the 'final model' (i.e. the model produced following adjustment of the BS model to account for optimism, as discussed in section 9.2.8). The regression coefficients and *p* values for the candidate predictors are given. The 'constant' for the model (i.e. the expected value of the log-odds of the outcome event when all predictor variables equal zero) is also reported; this permits calculation of the predicted risk for an individual patient using the standard logistic regression equation:

$$\text{Log-odds of the outcome event} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

where β_0 is the constant. The quantity on the right-hand side of the equal sign is the linear predictor of the log-odds of the outcome given the particular value of the p exposure variables x_1 to x_p (where x takes the numerical value of continuous variables, and is coded zero or one for categorical variables). The β's are the regression coefficients associated with the p exposure variables[181].

The predicted probability can then be calculated by exponentiating the log-odds (to give the odds of the outcome event), followed by use of the formula:

$$\text{Predicted probability} = \frac{odds}{(1 + odds)}$$

$$\text{Alternatively, predicted probability} = \frac{1}{(1 + \exp(-(log\ odds))}$$

9.2.6.3 Model diagnostics

The following model diagnostic checks were performed for the BS model:

- evidence for clustering caused by the duplicate admissions (to assess the suitability of random effects modelling);
- accuracy of the 'quadrature approximation' used to estimate the random effects model (to assess model reliability);
- evidence for 'specification error' (to assess whether the model included all relevant predictors);
- the impact of outlying observations (to identify the need to make adjustments to the model by removing or modifying influential observations).

Each is described below.

Evidence of clustering

After specifying a random effects model, Stata automatically reports the intra-cluster correlation (rho), which is a measure of the correlation within clusters. Where individuals within a cluster are no more similar than individuals in different clusters, rho is equal to zero. Where rho is one, everyone within a cluster acts the same. Stata also performs a likelihood ratio test, with the null hypothesis that there is no clustering (i.e. rho equal to zero).

Stata reports rho for models fitted with both non-imputed and multiply imputed data, but only reports the likelihood ratio test for non-imputed data. This is because the analysis of multiply imputed data involves an initial analysis being performed on each imputed dataset, with results then pooled using Rubin's combination rules. As a consequence, some postestimation procedures (such as the likelihood ratio test) are not directly applicable to multiply imputed data[176]. Rho was therefore reviewed after fitting the BS model in both the non-imputed and multiply imputed datasets, and the likelihood ratio test result reviewed after fitting the model using the non-imputed data.

Accuracy of the quadrature approximation

Random effects models are calculated using quadrature, which is an approximation whose accuracy depends partially on the number of integration points used[187]. After fitting a random effects model it is recommended that the accuracy of the quadrature approximation is checked[185 187]. With Stata, this can be performed using the 'quadchk' command[188]. This compares the parameter estimates obtained using different numbers of integration points, and calculates the relative difference in the regression

coefficients; if the coefficients do not change by more than a relative difference of 0.01%, this indicates that the choice of quadrature points does not significantly affect the outcome, and the results may be confidently interpreted[188].

The 'quadchk' command is not compatible with multiply imputed datasets; I therefore used the non-imputed data to check the accuracy of the quadrature approximation. This permitted inclusion of 1,494 (99.4%) of the 1,503 admissions, therefore providing a close approximation to the equivalent multiply imputed model in terms of the number of observations.

Testing for specification error

When building a logistic regression model, it is assumed that the log-odds of the outcome is a linear combination of the independent variables, and that all relevant independent variables have been included in the model. If this is not the case, the model is said to be 'miss-specified' (i.e. there is a specification error)[174].

After fitting a logistic regression it is possible to detect a specification error by refitting a model using the 'predicted value', and 'predicted value squared'. If the model is correctly specified, the 'predicted value' should be a statistically significant predictor (as it is the predicted value from the model). Conversely, the 'prediction squared' should have little predictive power except by chance. If the 'prediction squared' is significant, this suggests that relevant variables were excluded from the model, or that possible interactions have been overlooked[174].

The Stata command 'linktest' can be used to perform the above procedure in non-imputed data. As multiply imputed data were used to create the BS model, the specification error check was performed manually as follows:

1. The BS model was specified using the standard Stata commands, saving the estimation results in a file named 'miest';

```
. mi estimate, saving(miest,replace): xtlogit study_outcome num_comorb_F num_meds_F egfr_trun_dec_F wcc_trun_F allergy io(1 3
> 6 7).prim_diag  gent_vanc antimicrob epilepsy , re i(id)
```

2. 'Predicted value' (_hat) was generated using the saved estimation results;

```
. mi predict _hat using miest, xb
```

3. 'Prediction squared' (_hatsq) was generated, (i.e. '_hat' multiplied by '_hat');

```
. mi passive: gen _hatsq = _hat*_hat
```

4. A logistic regression model was specified using the 'predicted value', and 'predicted value squared'.

```
. mi estimate: logit study_outcome _hat _hatsq
```

It was then possible to review the statistical significance of '_hat' and '_hatsq', thereby testing for evidence of a specification error.

Detection and review of outlying observations

Outlying observations (cases) have the potential to yield biased regression coefficients as they may have significant leverage on the regression line. Following specification of the BS model I therefore checked for outliers and influential cases using the following residual and influence measures[174 189]:

- standardised Pearson residuals (a measure of the relative deviation between observed and fitted values);
- deviance residuals (a measure of disagreement between the observed and fitted log likelihoods functions);
- Pregibon leverage (a measure of the influence of observations).

The following diagnostic statistics were also used to identify cases with a substantial impact on the regression model in terms of:

- chi-square fit statistic (dx2);
- deviance statistic (dd);
- regression coefficients (Pregibon's dbeta).

The results were examined graphically by plotting each of the above measures against the model's predicted values.

The following standard rule of thumb cut-off values were used to identify outliers[174 189]:

- standardised Pearson or deviance residuals outside the range of plus or minus two;
- Pregibon leverage greater than three times the average leverage for the model;
- chi-square fit statistic or deviance statistic greater than 3.84 (i.e. the upper ninety-firth percentile of chi-square distribution with one degree of freedom).

A cut-off value of 0.04 was used for Pregibon's dbeta, selected after visually examining the plotted values.

The residuals, influence measure, and diagnostic statistics for each of the outliers were then tabulated. The impact of these outlying cases was assessed by fitting regression models with and without the outliers, and comparing the regression coefficients and *p* values for each predictor, and the overall model chi-square values.

The commands used by Stata to generate the residuals, influence measures, and diagnostic statistics are not compatible with random effects modelling, or multiply imputed datasets[190-192]. I therefore tested for influential cases after fitting a standard logistic regression model using non-imputed data. This excluded nine (0.6%) of the 1,503 admissions (due to missing data).

9.2.7 Assessing model performance

The two key measures of the performance of prognostic models are discrimination and calibration[94].

Discrimination refers to be ability of a prediction model to differentiate between those who do or do not experience the outcome event[94]. The c-index is the most commonly used performance measure to indicate the discriminatory ability of prognostic models, with the c-index being identical to the area under the Receiver Operating Characteristic (ROC) curve for models with binary oucomes[193]. The area under the ROC curve can be interpreted as the probability that a patient with the outcome is given a higher probability of the outcome by the model than a randomly chosen patient without the outcome; an uninformative model has an area of 0.5, and a perfect model has an area of one[70].

Calibration refers to the agreement between observed outcomes and predictions from the model, and can be assessed graphically by plotting predictions on the x-axis and the observed outcome frequency on the y-axis[194]. As the observed outcome is binary, smoothing techniques can be used to visualise the association, so displaying the direction and magnitude of model miscalibration across the probability range[94]. Perfect predictions should be on the 45 degree line, with an intercept of zero, and slope of one; deviation from this suggests imperfect calibration[194]. Alternatively, the observed proportion of outcome events for groups of patients with similar probabilities can be plotted, so comparing the mean predicted probability and mean observed outcome (although this is less precise[193]); for example, observed outcome by decile of prediction[193 194].

To assess the calibration of the BS model I therefore:

- reported the estimates for the calibration slope and intercept;
- plotted the predicted probability of an outcome event against the observed outcome frequency, using a 'locally weighted scatterplot smoothing' (lowess) line to visualise the association and permit identification of the direction and magnitude of any model miscalibration.

Discrimination was reported as the c-index.

Clinical usefulness, as assessed using decision curve analysis, is discussed in section 9.2.9.3.

9.2.8 Internal validation / adjustment for optimism

The predictive performance of prognostic models is overestimated when assessed using the same data used in development (known as the apparent performance), simply because the model has been optimised for that data. This results in overconfident predictions in independent data, where higher predictions are too high, and low predictions too low[195]. It is therefore recommended that all model development studies include some form of internal validation; for example split-sample validation (where the development data is divided into two datasets, one for model development and one for validation), or bootstrapping (which mimics the process of sampling from the underlying population by drawing random samples from the developmental dataset)[70 94]. Bootstrap validation is generally regarded as the preferred method as it permits all data to be used for model development, so is more statistically efficient[60]. In addition, bootstrap validation permits optimism to be quantified, and provides an estimate of any adjustments required[70 94].

Bootstrap validation involves the following steps[70 94 196]:

1. take a random bootstrap sample from the original dataset, identical in size, and drawn with replacement (therefore a patient may be included a number of times or not at all);
2. construct a model in the bootstrap sample, using similar steps to those used to develop the original regression model, then record the apparent performance of this bootstrap model (for example, c-index and calibration slope);
3. apply the bootstrap model to the original dataset[iv] and record the performance (known as the test performance);
4. calculate the optimism as the difference between bootstrap and test performance;
5. repeat steps 1 to 4 until stable estimates for the optimism are obtained (at least 100 times);
6. calculate the optimism for the c-index and calibration slope by averaging the results for each, as obtained in step 4.

Steyerberg advises that 100-200 bootstraps may be sufficient[70]; I chose to use 200 to increase the stability of the estimates.

[iv] The bootstrap model is applied to the original dataset using the regression equation for the model, as described in section 9.2.6.2. The regression equation is also known as the 'linear predictor' or 'prognostic index'.

The optimism calculated for the c-index was then used to correct the c-index of the BS model (by subtracting the value obtained in step 6). The c-index for the BS model prior to removal of 'non-sensible' predictors (as discussed in section 9.2.6.2) was used in this calculation. This was to permit direct comparison, that is to say, between models produced by backward elimination that contained all significant predictors (according to the predetermined 'stopping rule').

The optimism in the calibration slope was used to estimate the shrinkage factor required to adjust the model's regression coefficients to account for overconfident predictions. By definition, the average calibration slope of models developed at step 2 will be one, with an average intercept of zero; this is because calibration plots tend to show good calibration in the dataset from which they were developed[94]. When a model is used in new datasets (as in step 3) the intercept indicates the extent that predictions are systematically too high (intercept lower than zero) or too low (intercept higher than zero)[197]. In addition, a calibration slope of smaller than one suggests the regression coefficients in the original model were too large, resulting in predictions that are too extreme[197]. The average optimism for the calibration slope (calculated at step 6) can therefore be used as a 'linear shrinkage factor'. The initially estimated regression coefficients for the BS model were therefore adjusted for optimism by multiplying each by this shrinkage factor to create the 'final model' (i.e. the model used to create the MOAT).

Ideally the bootstrap models should be developed using the same modelling and predictor selection methods as used for the original model, but it is acknowledged that it is often difficult to replicate all steps in the bootstrap procedure[70]. I therefore chose to develop the bootstrap models using backward elimination, with the same significance level to exclude predictors as used for the original model (*p* greater than 0.157). Manual predictor selection was used for the original model, but for pragmatic reasons (related to time required to perform manual backward elimination in each of the 200 bootstrap samples) I chose to use an automated selection technique for the bootstrap models. As it is not possible to use Stata's automated selection techniques with either multiply imputed data, or random effects models[198], I chose to perform the bootstrap validation using standard logistic regression, and non-imputed data. I anticipated this would have minimal impact, as the purpose of bootstrap validation is to quantify possible optimism rather than to provide accurate estimates of the regression coefficients. To minimise the impact of missing data, I excluded BMI from the full model used to develop the bootstrap models (due to the large number of missing data for this

variable). I anticipated this would have minimal impact on the bootstrap models as both univariable and multivariable analyses suggested no evidence that BMI was associated with the outcome event. This permitted inclusion of 1,447 (96.3%) of the 1,503 admissions in the bootstrap validation.

9.2.9 Development of a decision aid (the MOAT)

Development of the MOAT (a decision aid to permit targeting of patients at high risk of experiencing MSP MRPs) involved three stages:

- selection of a 'presentation format' (for example a simplified paper score chart, or electronic scoring system);
- creation of risk groups;
- assessment of the potential clinical usefulness of the MOAT (based on the risk groups chosen).

Each of these is described below.

9.2.9.1 Presentation format

Prognostic modelling studies often aim to develop a simplified scoring system to permit use of the model in clinical practice without the use if complex calculations (for example by simplifying regression equations to create an easy-to-sum score)[94]. Although this was my original intention, as stated in the MOAT study protocol[142], I subsequently decided to use an 'electronic' format for the MOAT (permitting use of the final model's full regression equation for risk predictions). The reasons for this were:

- even with a simplified score chart, pharmacy staff would be required to assign 'points' to each predictor, then sum the final score. Use of an electronic tool simply requires pharmacy staff to input the raw predictor data, then calculations are performed automatically, so reducing the time needed to apply the MOAT, and the possibility of calculation errors;
- simplification of prognostic models generally leads to loss of predictive accuracy (due to rounding of regression coefficients)[94]. This would be avoided by use of an electronic format;
- there is an aim for all English hospitals to be paperless by 2020[199], which has resulted in the increased use of laptop computers by pharmacy staff at ward level, increasing routine access to electronic resources;
- usage instructions for the MOAT can be incorporated into an electronic tool using 'help' links, so increasing accessibility to guidance. I anticipated this may increase consistency in use of the MOAT, and reduce barriers to use (related to lack of familiarity, or confusion over which data are required).

The regression equation for the final model was therefore used to create an electronic scoring system. Microsoft Excel was used due to its wide availability. Programming

was carried out by Aneesh Khurana (IT Systems Manager for Pharmacy at the Luton and Dunstable University Hospital).

9.2.9.2 Creation of risk groups

While prognostic models provide estimates of the probability that an individual patient will experience an outcome event, this does not provide guidance to users of the model on an appropriate course of action. 'Risk groups' are therefore often created, which indicate a specific course of action, so creating a 'decision aid' or 'clinical decision rule'[70]. No consensus exists on how risk groups should be created[94], but it is recommended that subject matter input is used rather than reliance on statistical estimation[70]. I therefore sought advice from a group of practising pharmacy staff (as part of the assessment of the MOAT, discussed in chapter 10) to inform the choice of cut-offs for the predicted risk probabilities (also known as decision thresholds or classification cut-offs)[70 194]. Concern exists over the arbitrary nature of categorisation, with all patients within a group being assumed to have the same risk[94]. I therefore chose to create risk groups, but to report both the risk group and individual predicted risk probability for each patient assessed using the MOAT. This was to guide general prioritisation decisions (by categorising patients as high, medium or low-risk), but also to permit some degree of prioritisation within each category (if required due to workload pressures).

The predictive performance of the MOAT is reported using the classification measures: sensitivity and specificity, which are reported for the decision thresholds chosen for the medium and high-risk groups. The predicted risk probability range for each of the risk groups is also reported.

9.2.9.3 Assessment of clinical usefulness

Once decision thresholds were selected, it was possible to assess clinical usefulness, in terms of whether the MOAT is likely to be beneficial in clinical practice for guiding decision making[70]. This goes beyond calculation of the c-index (which is primarily interested in predictive accuracy) to incorporate information on consequences, for example considering the relative impact of false negative and false positive predictions[200].

Decision curve analysis has been suggested as a method to assess clinical usefulness[70 193]. This permits performance of a model to be assessed over a range of decision thresholds, using the theoretical relationship between threshold probabilities

and the relative value of false positive and false negative results[200], calculated as the net benefit:

$$Net\ benefit = \frac{true\ positive\ count}{total\ number\ of\ patients} - \frac{false\ postive\ count}{total\ number\ of\ patients} \times \left(\frac{threshold\ probability}{1 - threshold\ probability}\right)$$

Net benefit is interpreted in units of the true positives, and is a measure of how many more patients are correctly 'treated' (true positives) at the same rate of 'not treating' those who do not need treatment (false positives)[70].

By varying the threshold probability it is possible to produce a 'decision curve', with threshold probability plotted on the x-axis, and net benefit plotted on the y-axis. The net benefit of default policies of 'treat none' and 'treat all' are also plotted to permit comparisons to be made. The net benefit of 'treating none' is zero (as the true and false positive counts are both zero), therefore if the net benefit of the prediction model is positive, it is better to use the model than 'treat none'. The true and false positive counts for the 'treat all' strategy are the number of patients with and without the outcome respectively; the net benefit of 'treat all' is therefore equal to the outcome prevalence at a threshold probability of zero, and equal to zero at the prevalence of the outcome[200].

The decision curve informs the range of threshold probabilities for which the prediction model would be of value in clinical practice[201].To interpret a decision curve one identifies a range of plausible threshold probabilities, and then determines whether the model has benefits (i.e. a net benefit greater than 'treating all' and 'treating none') at all values within this range[70]. I therefore plotted the decision curve for the MOAT to establish whether it has the potential to be clinically useful at the decision thresholds selected in section 9.2.9.2.

9.3 Results

The results are presented according to the TRIPOD reporting guidelines[57]. This section includes:

- exploratory data analysis;
- imputation of missing data;
- model development;
- assessment of model performance;
- internal validation / adjustment for model 'optimism';
- development of the MOAT (decision aid based on the regression model).

Each of these will be described in turn.

9.3.1 Exploratory data analysis

9.3.1.1 Review of distribution of categorical predictors

A review of the distributions of categorical candidate predictors pre-selected for MOAT development identified five categories containing fewer than 5% of the study population (Table 17, chapter 6). All were categories of 'high-risk medicines':

- theophylline and aminophylline (2.5% of study population)
- immunosuppressants (1.4%)
- cytotoxics (0.9%)
- lithium (0.4%)
- 'other high-risk medicines', which included clozapine, anti-retrovirals, and medicines for Parkinson's disease (2.7%)

Given the low prevalence of each of the above categories I decided to combine them, so creating a larger 'other high-risk medicines' category. I felt this was clinically sensible, as it prevented any of the high-risk medicines being excluded from the analyses. In addition, the 'other' category already contained medicines with diverse pharmacological uses; as discussed in chapter 6, the overall categorisation of high-risk medicines was recognised as simplistic, albeit necessary to prevent model overfitting. I also moved clozapine from 'other high-risk medicines' to 'antipsychotics'. This was on the basis that clozapine was used very infrequently (in only one of the 1,503 study admissions), in addition to clozapine being more closely related to the remaining antipsychotics in terms of pharmacological use. This resulted in the 'other high-risk category' increasing from 2.7% of the sample population to 7.8%, so permitting robust

modelling. The prevalence of ‘antipsychotics’ increased marginally (from 6.1% to 6.2%).

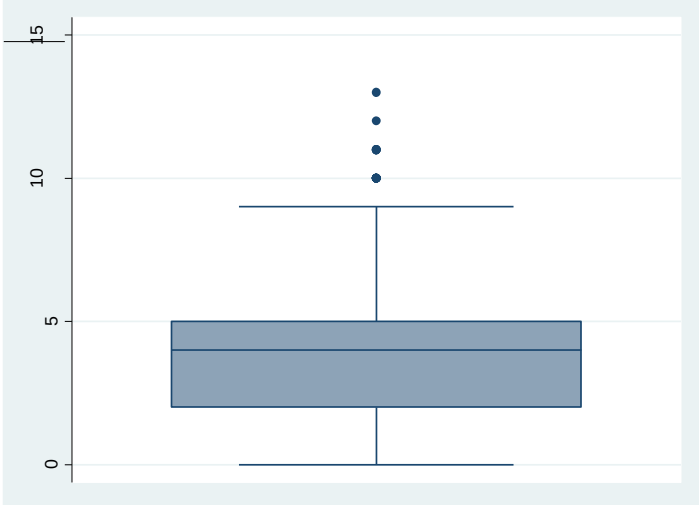
The high-risk medicine categories included in model development (following the above review) are listed below:

- anticoagulants / direct oral anticoagulants;
- therapeutic heparin;
- anti-diabetic medication;
- opiates (excluding codeine, tramadol, meptazinol and dihydrocodeine);
- systemic aminoglycosides and glycopeptides;
- systemic antimicrobials (excluding aminoglycosides and glycopeptides);
- epilepsy medicines;
- antipsychotics;
- antiarrhythmics;
- antidepressants;
- other high-risk medicines (including anti-retrovirals, cytotoxics, immunosuppressants, lithium, medicines for Parkinson’s disease).

Details of the specific medicines included in each category (based on the MOAT sample population) are given in Appendix A9.1.

9.3.1.2 Identification and review of outliers (continuous predictors)

The 12 continuous candidate predictors pre-selected for MOAT development were assessed for the presence of outlying values using box-plots and tables of extreme values. Figure 4 and Figure 5 provide examples of the respective Stata outputs.



Central box represents interquartile range, with median shown as central line.

Whiskers show the 'adjacent line' (one and a half times the interquartile range above or below the third or first quartile respectively).

Outlying values shown as points above or below the adjacent line.

Figure 4 – Box-plot of variable 'number of comorbidities'

.

obs:	iqr:	num_co~b
45.	1.667	10
231.	1.667	10
253.	1.667	10
332.	1.667	10
653.	1.667	10
826.	1.667	10
861.	1.667	10
949.	1.667	10
997.	1.667	10
1216.	1.667	10
291.	2.000	11
818.	2.000	11
867.	2.000	11
1047.	2.000	11
516.	2.333	12
226.	2.667	13

obs = observation number (i.e. unique identifier)

iqr = number of interquartile ranges above or below third or first quartile respectively

num_comb = variable name for 'number of comorbidities'

Figure 5 – Stata output for 'extremes' module

Table 38 gives the number (and percentage) of outliers for each predictor, with the range of values. The number of outliers per variable ranged from one data point (0.07%) for 'age', to 132 (8.8%) for 'number of hospital admissions in previous six months'.

Table 38 – Number and ranges for outliers

Predictor	Outlier(s)* (admissions = 1,503) n (% of admissions†)		Value / range of outlier(s)	
	Below lower truncation point	Above upper truncation point	Below lower truncation point	Above upper truncation point
Age (years)	1 (0.07)	0	17	N/A
Socioeconomic status, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]	0	0	N/A	N/A
Body mass index (kg/m^2)	0	30 (2.6)	N/A	40.8-65.5
Number of hospital admissions in previous 6 months	0	132 (8.8)	N/A	3-10
Number of comorbidities	0	16 (1.1)	N/A	10-13
Number of medicines prescribed	0	26 (1.7)	N/A	18-27
Renal function - estimated glomerular filtration rate‡ (ml/min/1.73m^2)	0	38 (2.5)	N/A	162-309
Serum albumin (g/L)	34 (2.3)	5 (0.3)	7-19	48-55
Serum potassium (mmol/L)	9 (0.6)	18 (1.2)	2.3-2.8	6-7.8
Serum sodium (mmol/L)	68 (4.5)	20 (1.3)	111-127	148-170
White cell count (10^9/L)	0	64 (4.3)	N/A	20.8-93.0
Platelet count (10^9/L)	3 (0.2)	51 (3.4)	5-7	492-977

N/A = not applicable

* Limit set at one and a half times the interquartile range above or below the third or first quartile respectively

† Calculated as the number of outliers / number of available values for each variable (i.e. excludes admissions with missing data)

‡ Glomerular filtration rate estimated using modified Modification of Diet in Renal Disease (MDRD) equation[121]

Plausibility of the outlying values was assessed by reviewing the distribution of each predictor across the full range of observed values. This found that all outlying results were clinically plausible, and consistent with the overall data distributions, that is, no data points were significantly remote, with outliers representing extreme values within skewed or wide distributions (illustrated by Figure 6 and Figure 7 respectively). No obvious data entry errors were identified; therefore no data were set to missing.

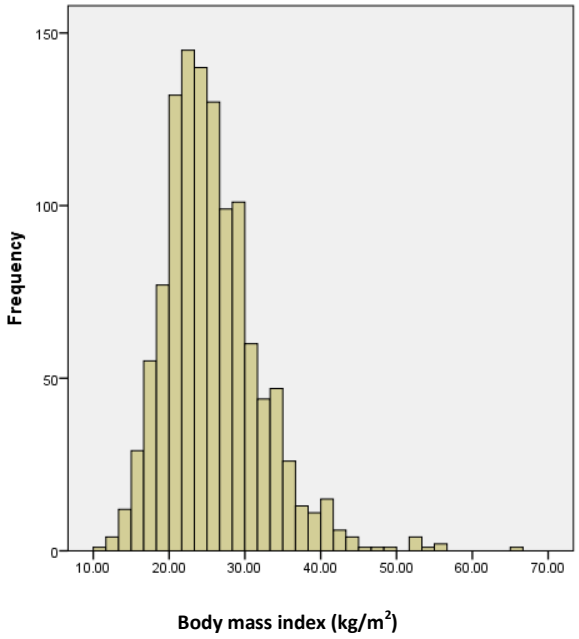


Figure 6 – Histogram showing distribution of body mass index

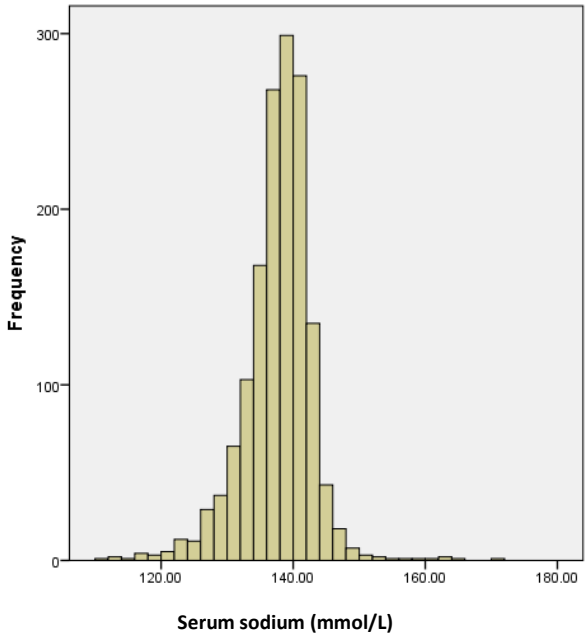


Figure 7 – Histogram showing distribution of serum sodium

The results of univariable logistic regressions using non-truncated and truncated data are given in Table 39. The impact of truncation, in terms of the difference in the regression coefficients for individual predictors, ranged from zero to 170%, with a 'substantial difference' (i.e. a change of 10% or more) observed for seven of the 12 predictors (suggesting that the affected predictors had outlying data points that were not representative of the remaining sample population). For example, truncation of 64 outlying results for 'white cell count', representing 4.3% of admissions, caused an increase in the regression coefficient of 125%; inclusion of these outlying data points in the analysis therefore had the potential to significantly alter the predictor-outcome relationship.

Table 39 also shows that truncation resulted in an increased model chi-square for most predictors, suggesting that outlying values generally caused underestimation of predictive effect. This was not the case for 'number of comorbidities', 'serum albumin' (upper truncation), 'serum potassium' (upper truncation), and 'serum sodium' (lower truncation), where outlying values appear to overestimate the predictive effect. While this is only indicative of the potential impact of truncation on the statistical significance of the univariable models, it provides some justification for the method chosen to decide whether or not to truncate outlying data. That is, basing the decision on the percentage change in regression coefficients is an objective approach, driven by a desire to obtain accurate parameter estimates, rather than simply maximise model fit, which could lead to overoptimistic predictions[171].

Regarding the clinical sensibility of the truncation points, the upper truncation point for albumin was lower than the standard reference range (i.e. 50 g/L). All other truncation points were outside standard ranges. As shown in Table 39, it was not necessary to truncate the high albumin values, therefore no further action was needed to address this anomaly.

Table 39 – Univariable regression of truncated and non-truncated data

Predictor and truncation point	Regression coefficient*			Truncation selected for model development	Model chi-square†	
	Non-truncated model	Truncated model‡	Percentage change		Non-truncated model	Truncated model‡
Age/10 (years) Lower truncation point = 18 years	0.167	0.167	0	No	33.87	33.86
Body mass index (kg/m^2) Upper truncation point = 40.6 kg/m^2	0.0042	0.0052	23.8	Yes	0.20	0.27
Number of hospital admissions in previous 6 months Upper truncation point = 2	0.0643	0.1735	169.8	Yes	2.48	6.96
Number of comorbidities Upper truncation point = 9	0.2131	0.2164	1.6	No	82.34	81.91
Number of medicines prescribed Upper truncation point = 17	0.0952	0.0992	4.2	No	53.00	53.64
Estimated glomerular filtration rate§/10 (ml/min/1.73m^2) Upper truncation point = 161 ml/min/1.73m^2	-0.079	-0.088	11.4	Yes	28.82	30.77
Serum albumin (g/L) Upper truncation point = 47 g/L	-0.0279	-0.0278	0.4	No	9.86	9.70
Serum albumin (g/L) Lower truncation = 20 g/L		-0.0312	11.8	Yes		11.46
Serum potassium (mmol/L) Upper truncation point = 5.9 mmol/L	0.0964	0.0821	14.8	Yes	1.28	0.85
Serum potassium (mmol/L) Lower truncation = 2.9 mmol/L		0.1017	5.5	No		1.40
Serum sodium (mmol/L) Upper truncation point = 147 mmol/L	-0.0114	-0.0123	7.9	No	1.29	1.34
Serum sodium (mmol/L) Lower truncation point = 128 mmol/L		-0.0109	4.4	No		0.93
White cell count (10^9/L) Upper truncation point = 20.7 10^9/L	0.0126	0.0283	124.6	Yes	2.28	5.30
Platelet count (10^9/L) Upper truncation point = 490 10^9/L	0.0020	0.0023	15	Yes	0.16	0.17
Platelet count/10 (10^9/L) Lower truncation point =14 10^9/L		0.0020	0	No		0.16

* Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

† Based on the likelihood ration chi-square test (difference in the log likelihood of a model with and without the independent variable)

‡ Truncation points set at one and a half times the interquartile range above or below the third or first quartile respectively

§ Glomerular filtration rate estimated using modified Modification of Diet in Renal disease (MDRD) equation[121]

9.3.1.3 Linearity

Figure 8 shows the Stata output for the full MFP model (i.e. including all predictors). As MFP is performed in cycles, the output gives the results for the ‘initial’ and ‘final’ cycles, with the ‘final’ cycle showing the powers selected for the MFP model. The output shows that MFP modelling selected a power of ‘1’ for all predictors (i.e. no transformation required).

Final multivariable fractional polynomial model for study_outcom

Variable	Initial			Final		
	df	Select	Alpha	Status	df	Powers
age_dec_F	4	1.0000	0.0500	in	1	1
ses_dec_F	4	1.0000	0.0500	in	1	1
bmi_trun_F	4	1.0000	0.0500	in	1	1
num_hosp_...	1	1.0000	0.0500	in	1	1
num_comor...	4	1.0000	0.0500	in	1	1
num_meds_F	4	1.0000	0.0500	in	1	1
egfr_trun_F	4	1.0000	0.0500	in	1	1
albumin_t...	4	1.0000	0.0500	in	1	1
potassium...	4	1.0000	0.0500	in	1	1
sodium_F	4	1.0000	0.0500	in	1	1
wcc_trun_F	4	1.0000	0.0500	in	1	1
platelets...	4	1.0000	0.0500	in	1	1
allergy	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
_Iprim_di...	1	1.0000	0.0500	in	1	1
dementia	1	1.0000	0.0500	in	1	1
iv_use	1	1.0000	0.0500	in	1	1
anticoags	1	1.0000	0.0500	in	1	1
heparin	1	1.0000	0.0500	in	1	1
diabetes	1	1.0000	0.0500	in	1	1
opiates	1	1.0000	0.0500	in	1	1
gent_vanc	1	1.0000	0.0500	in	1	1
antimicrob	1	1.0000	0.0500	in	1	1
epilepsy	1	1.0000	0.0500	in	1	1
antipsychot	1	1.0000	0.0500	in	1	1
antiarryth	1	1.0000	0.0500	in	1	1
antidepres	1	1.0000	0.0500	in	1	1
other_hig...	1	1.0000	0.0500	in	1	1
liver_dx	1	1.0000	0.0500	in	1	1

df = degrees of freedom, which is twice the degree of the fractional polynomial (FP). The df for the 'initial' model represents the FPs considered in the first cycle of multivariable fractional polynomial (MFP) modelling, and the df for the 'final' model represents the FP selected for the final model

Select = nominal significance level for variable selection by backward elimination (set at '1' to force all variables to remain in the model (i.e. full model produced)

Alpha = significance levels for testing between different FP models (0.05 is the default nominal *p* value)

Status = whether variable selected during backward elimination (all variables 'in' as backwards elimination not used)

Power = FP power selected from the default set (−2, −1, −0.5, 0, 0.5, 1, 2, 3), where inverse is x^{-1}, square root is $x^{0.5}$, log is x^{0}, linear is x^{1}, squared is x^{2}, and so on

Figure 8 – Stata output for full multivariable fractional polynomial model

Figure 9 shows the Stata output for the MFP model produced using backward elimination. This shows that linearity was not altered following removal of non-significant variables. It should be noted that this backward elimination was performed using complete-case analysis (i.e. prior to multiple imputation). Results may therefore differ following multiple imputation (as discussed in section 9.2.5).

Final multivariable fractional polynomial model for study_outcome

Variable	Initial			Final		
	df	Select	Alpha	Status	df	Powers
age_dec_F	4	0.1570	0.0500	out	0	
ses_dec_F	4	0.1570	0.0500	out	0	
bmi_trun_F	4	0.1570	0.0500	out	0	
num_hosp_...	1	0.1570	0.0500	out	0	
num_comor...	4	0.1570	0.0500	in	1	1
num_meds_F	4	0.1570	0.0500	out	0	
egfr_trun_F	4	0.1570	0.0500	out	0	
albumin_t...	4	0.1570	0.0500	out	0	
potassium...	4	0.1570	0.0500	out	0	
sodium_F	4	0.1570	0.0500	out	0	
wcc_trun_F	4	0.1570	0.0500	in	1	1
platelets...	4	0.1570	0.0500	out	0	
allergy	1	0.1570	0.0500	in	1	1
_Iprim_di...	1	0.1570	0.0500	in	1	1
_Iprim_di...	1	0.1570	0.0500	out	0	
_Iprim_di...	1	0.1570	0.0500	in	1	1
_Iprim_di...	1	0.1570	0.0500	in	1	1
_Iprim_di...	1	0.1570	0.0500	out	0	
_Iprim_di...	1	0.1570	0.0500	out	0	
_Iprim_di...	1	0.1570	0.0500	out	0	
dementia	1	0.1570	0.0500	out	0	
iv_use	1	0.1570	0.0500	out	0	
anticoags	1	0.1570	0.0500	out	0	
heparin	1	0.1570	0.0500	in	1	1
diabetes	1	0.1570	0.0500	out	0	
opiates	1	0.1570	0.0500	out	0	
gent_vanc	1	0.1570	0.0500	in	1	1
antimicrob	1	0.1570	0.0500	in	1	1
epilepsy	1	0.1570	0.0500	in	1	1
antipsychot	1	0.1570	0.0500	out	0	
antiarryth	1	0.1570	0.0500	out	0	
antidepres	1	0.1570	0.0500	in	1	1
other_hig...	1	0.1570	0.0500	out	0	
liver_dx	1	0.1570	0.0500	out	0	

See Figure 8 for explanation of abbreviations

Figure 9 – Stata output for multivariable fractional polynomial model selected using backward elimination

Both models therefore failed to reject linear relationships for the continuous predictors (given the default ‘alpha’ of 0.05). It was therefore not necessary to include non-linear transformations in the imputation model, or during model building.

9.3.1.4 Multicollinearity

The VIF for the candidate predictors is shown in Table 40. The VIFs ranged from 1.11 (for 'socioeconomic status', and 'other high-risk medicines') to 4.67 (for the primary diagnosis 'respiratory'). As all VIFs were below 10, this suggests no excessive correlation between predictors. As a result, it was not necessary to take further action regarding multicollinearity.

Table 40 – Variance inflation factors for the candidate predictors

Predictor	Variance inflation factor
Age	1.94
Socioeconomic status	1.11
Previous allergy / adverse drug reaction	1.13
Body mass index	1.24
Number of hospital admissions in previous 6 months	1.20
Number of comorbidities	1.82
History of dementia	1.20
Number of medicines prescribed	1.64
Parenteral administration route	1.48
Estimated glomerular filtration rate	1.50
Liver disease	1.16
Serum albumin	1.38
Serum potassium	1.19
Serum sodium	1.17
White cell count	1.36
Platelet count	1.32
High-risk medicines:	
Anticoagulants	1.33
Therapeutic heparin	1.36
Anti-diabetic medication	1.25
Opiates	1.13
Aminoglycosides and glycopeptides	1.14
Other antimicrobials	1.61
Epilepsy medicines	1.18
Antipsychotics	1.08
Antiarrhythmics	1.27
Antidepressants	1.13
Other high-risk medicines	1.11
Primary diagnosis:	
Nervous system and mental disorders	2.90
Cardiovascular system	4.62
Respiratory system	4.67
Gastrointestinal system	2.67
Genitourinary system	2.94
Musculoskeletal-intergumentary systems	2.32
All other categories of primary diagnosis	3.66

9.3.1.5 Univariable analyses

The results of the univariable analyses are shown in Table 41. There was strong evidence for statistically significant associations between the outcome event and the following predictors (p <0.05):

- age;
- socioeconomic status;
- previous allergy;
- number of hospital admissions in previous six months;
- primary diagnosis (as a combined effect);
- number of comorbidities;
- estimated glomerular filtration rate;
- serum albumin;
- white cell count;
- number of medicines prescribed;
- use of anticoagulants, anti-diabetic medication, aminoglycosides and glycopeptides, other antimicrobials, epilepsy medicines, and antidepressants;
- parenteral medicine administration.

There was also weak evidence for an association between the outcome and the use of therapeutic heparin (p = 0.057), opiates (p = 0.073), and antiarrhythmics (p = 0.053). There was no evidence for an association between the remaining candidate predictors and the outcome event.

Table 41 – Univariable association between predictors and outcome events

Predictor	Occurrence of outcome event		Odds ratio‡ (95% CI)	Univariable *p* value§
	No (%)	Yes (%)		
Demographic				
Age/10 (years)	Mean: 6.78 (i.e. 67.8 years)	Mean: 7.35 (i.e. 73.5 years)	1.18 (1.12 to 1.25)	<0.001
Socioeconomic status/10*, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]†	Mean: 5.2	Mean: 5.5	1.04 (1.00 to 1.08)	0.044
Patient related				
Previous allergy*	302 (33.8)	280 (46.0)	1.67 (1.35 to 2.06)	<0.001
Body mass index* (kg/m^2)	Mean: 25.6	Mean: 25.7	1.01 (0.99 to 1.03)	0.610
Number of hospital admissions in previous 6 months	Mean: 0.56	Mean: 0.67	1.19 (1.05 to 1.35)	0.008
Primary diagnosis:				Combined effect 0.0078
Endocrine and metabolic	46 (5.2)	36 (5.9)	1.16 (0.74 to 1.81)	0.530
Nervous system and mental disorders	82 (9.2)	67 (11.0)	1.22 (0.87 to 1.72)	0.252
Cardiovascular system	186 (20.8)	129 (21.2)	1.02 (0.79 to 1.31)	0.882
Respiratory system	197 (22.1)	135 (22.1)	1.00 (0.78 to 1.29)	0.974
Gastrointestinal system	105 (11.8)	39 (6.4)	0.51 (0.35 to 0.75)	<0.001
Genitourinary system	76 (8.5)	68 (11.2)	1.35 (0.96 to 1.90)	0.090
Musculoskeletal-intergumentary systems	47 (5.3)	46 (7.5)	1.47 (0.96 to 2.23)	0.074
All other categories	154 (17.3)	90 (14.8)	0.83 (0.63 to 1.10)	0.197
Number of comorbidities	Mean: 3.3	Mean: 4.4	1.24 (1.18 to 1.30)	<0.001
History of dementia	94 (10.5)	67 (11.0)	1.05 (0.75 to 1.46)	0.779
Laboratory results				
Estimated glomerular filtration rate/10* (ml/min/1.73m^2)	Mean: 8.03 (i.e. 80.3 ml/min/1.73m^2)	Mean: 7.04 (i.e. 70.4 ml/min/1.73m^2)	0.92 (0.89 to 0.95)	<0.001
Liver disease	101 (11.3)	63 (10.3)	0.90 (0.65 to 1.26)	0.548
Serum albumin* (g/L)	Mean: 33.5	Mean: 32.4	0.97 (0.95 to 0.99)	<0.001
Serum potassium* (mmol/L)	Mean: 4.4	Mean: 4.4	1.09 (0.91 to 1.29)	0.358
Serum sodium* (mmol/L)	Mean: 137.3	Mean: 137.0	0.99 (0.97 to 1.01)	0.257
White cell count* (10^9/L)	Mean: 10.4	Mean: 10.9	1.03 (1.00 to 1.05)	0.021

Continued from previous page…

Predictor	Occurrence of outcome event		Odds ratio‡ (95% CI)	Univariable *p* value§
	No (%)	Yes (%)		
Platelet count/10* (10^9/L)	Mean: 25.59 (i.e. 255.9 10^9/L)	Mean: 25.80 (i.e. 258.0 10^9/L)	1.00 (0.99 to 1.01)	0.682
Medicines related				
Number of medicines	Mean: 7.2	Mean: 8.7	1.10 (1.07 to 1.13)	<0.001
Use of high-risk medicines:				
Anticoagulants	161 (18.0)	151 (24.8)	1.50 (1.16 to 1.92)	0.002
Therapeutic heparin	119 (13.3)	103 (16.9)	1.32 (0.99 to 1.76)	0.057
Anti-diabetic medication	146 (16.4)	153 (25.1)	1.71 (1.33 to 2.21)	<0.001
Opiates	76 (8.5)	69 (11.3)	1.37 (0.97 to 1.93)	0.073
Aminoglycosides and glycopeptides	49 (5.5)	56 (9.2)	1.74 (1.17 to 2.59)	0.006
Other antimicrobials	512 (57.3)	525 (69.7)	1.71 (1.38 to 2.13)	<0.001
Epilepsy medicines	104 (11.7)	123 (20.2)	1.92 (1.44 to 2.55)	<0.001
Antipsychotics	50 (5.6)	43 (7.1)	1.28 (0.84 to 1.95)	0.255
Antiarrhythmics	78 (8.7)	72 (11.8)	1.40 (1.00 to 1.96)	0.053
Antidepressants	182 (20.4)	169 (27.7)	1.50 (1.18 to 1.90)	0.001
Other high-risk medicines	62 (6.9)	55 (9.0)	1.33 (0.91 to 1.94)	0.143
Parenteral administration	569 (63.7)	439 (72.0)	1.46 (1.17 to 1.83)	<0.001

* For patients without missing data (further details provided in Appendix A6.6)

† Deprivation rank based on patients' postcode, shown as the ranked position as a percentage of all neighbourhoods in England (where 1 is the most deprived)

‡ Measure of association between exposure and outcome event (the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of exposure)

§ Test for difference between admissions with and without occurrence of outcome event. Obtained from univariable regression modelling (based on the likelihood ratio chi-square test)

Age, socioeconomic status, estimated glomerular filtration rate, and platelet count analysed as deciles to aid interpretability (i.e. after dividing the actual value by ten)

9.3.2 Missing data

Chapter 6 identified the proportion of study admissions with missing data, with possible reasons for the missingness. Of the 1,503 admissions included in the regression analyses, 449 (29.9%) had one or more missing data point (Table 18). Excluding ethnicity, (as it is not a candidate predictor) the total number of admissions with one or more missing data point was 387 (25.7%). The number of missing data points for each candidate predictor is shown in Table 42. This shows that 430 data points (i.e. 1.6% of the total candidate predictor data) were missing.

Table 42 – Details of missing data for candidate predictors

Candidate predictor	**Admissions with missing data (admissions = 1,503)** n (% of admissions)
Age	0
Socioeconomic status	6 (0.4)
Previous allergy	1 (0.07)
Body mass index	341 (22.7)
Number of hospital admissions in previous 6 months	0
Primary diagnosis	0
Number of comorbidities	0
History of dementia	0
Number of medicines	0
High-risk medicines use	0
Parenteral administration route	0
Renal function	9 (0.6)
Liver disease	0
Serum albumin	26 (1.7)
Serum potassium	30 (2.0)
Serum sodium	3 (0.2)
White cell count	6 (0.4)
Platelet count	8 (0.5)
Total missing data points n (% for all predictors i.e. 27,054*)	430 (1.6)

* Total number of data points for all predictors calculated as the total number of predictors multiplied by the total number of study admissions

The remaining results for the missing data analysis are presented in five sections: use of common sense approaches to predict missing data, analysis of data distribution prior to imputation, comparison of observed and imputed values, imputation diagnostics, and the sensitivity analysis. Each is described below.

9.3.2.1 Use of common sense approaches to predict missing data values

Exploration of the missing data identified nine missing values that could be predicted using common sense solutions[182]:

- allergy status was undetermined for one admission. As this was a binary variable (yes/no), and the patient was prescribed nine medicines during their admission (including antimicrobials), I decided to count this admission as 'no allergies' for the purposes of the study (as this was the most likely result given that nursing staff at the study sites are required to check a patient's allergy status before administering any medication);
- three admissions (from Hospital B) had missing serum creatinine values, but estimated glomerular filtration rates (eGFR) were reported, calculated using the enzymatic method (potentially due to their high serum bilirubin levels, as this interferences with the standard creatinine assay[202]). Of the three admissions, one had an actual eGFR value reported, but two had the result reported as '>90 ml/min/1.73m^2' (a cut-point used to represent 'normal' renal function[125]). As it was not possible to obtain an actual eGFR value for these two patients, I chose to use 126 ml/min/1.73m^2, as this is mid-way between 91 ml/min/1.73m^2, representing the minimum possible value, and the eGFR upper truncation point of 161 ml/min/1.73m^2, which represented the highest value eGFR used for the present study;
- two admissions (Hospital A) with missing BMI data were found to be duplicate admissions (less than one month apart), therefore BMI results for the duplicate admission were used;
- Socioeconomic status was unavailable for six admissions because the method used to determine socioeconomic status was based on English postcodes, and two participants had no fixed abode, and four were not resident in England. Of the participants not resident in England, one lived in Scotland, and the remaining three were from Canada, Africa and Ibiza. It was possible to use the Scottish Index of Multiple Deprivation 2016[203] to establish a deprivation rank for the Scottish participant; while not directly comparable with the English rankings, it did provide an informed estimate. For the two participants with no fixed abode, I chose to rank the individuals in the lowest socioeconomic group (i.e. an Index of Multiple Deprivation Rank of one). This was on the basis that the indices are mainly informed by income and employment (total 45% of weighting), with additional

contributions from living environment (9.3%), crime (9.3%), and barriers to housing services (9.3%), all of which are likely to be impacted by homelessness.

Inclusion of this data reduced the number of admission with one or more missing data point to 379 (25.2%) of 1,503. It was therefore possible to include 1,124 complete cases in the multiple imputation sensitivity analysis (Table 43).

9.3.2.2 Analysis of data distribution

Review of the distributions of the continuous candidate predictors found that data were compatible with normal distributions (following the truncation of outlying data, as discussed in 9.3.1.2). Of the continuous auxiliary variables, 'weight' and 'height' appeared to follow normal distributions, but 'length of hospital stay' was positively skewed, as shown in Figure 10, with a skewness statistic of 4.0.

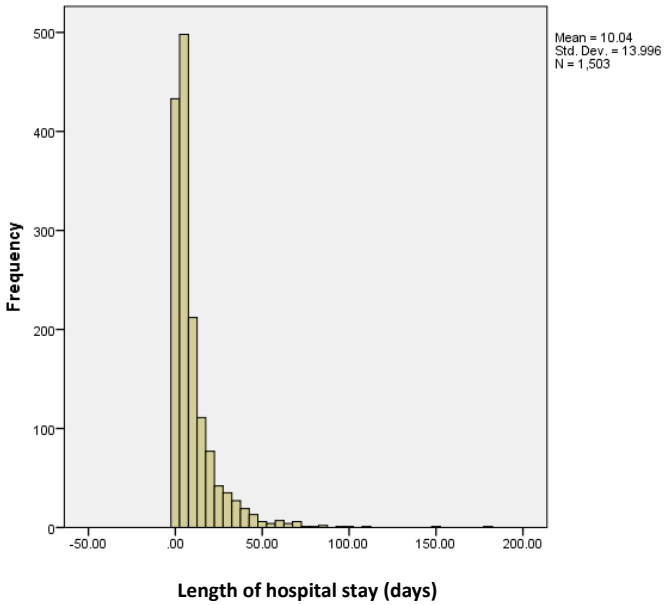


Figure 10 – Histogram showing distribution of length of hospital stay

Following use of a log transformation the skewness statistic reduced to 0.346, therefore a log transformation for 'length of hospital stay' was used for the imputation model.

9.3.2.3 Comparison of observed and imputed values

A review of the imputed values produced by the initial MICE imputation (specified using linear regression as the conditional distribution for the continuous variables) identified implausible values for: BMI, socioeconomic status, estimated glomerular filtration rate, and white cell count.

The imputation model was then specified using truncated distributions for the affected variables:

```
. mi impute chained (truncreg, ll(10.6)) bmi_trun_F (truncreg, ll(.3)) egfr_trun_dec_F (truncreg, ll(.3) ul(10)) ses_dec_F (t
> runcreg, ll(0.3)) wcc_trun_F (truncreg) albumin_trun_F potassium_trun_F sodium_F platelets_trun_dec_F weight height = age_d
> ec_F num_hosp_adm_trun_F num_comorb_F num_meds_F i.prim_diag allergy dementia liver_dx iv_use anticoags heparin diabetes op
> iates gent_vanc antimicrob epilepsy antipsychot antiarryth antidepres other_high_risk study_outcome ln_length_stay male , a
> dd (20) rseed (53421) savetrace(trace1,replace)
```

where:

- ‘trunreg’ permits specification of the upper and/or lower limits for the distributions of individual variables;
- ‘ll’ sets the lower limit, and ‘ul’ sets the upper limit.

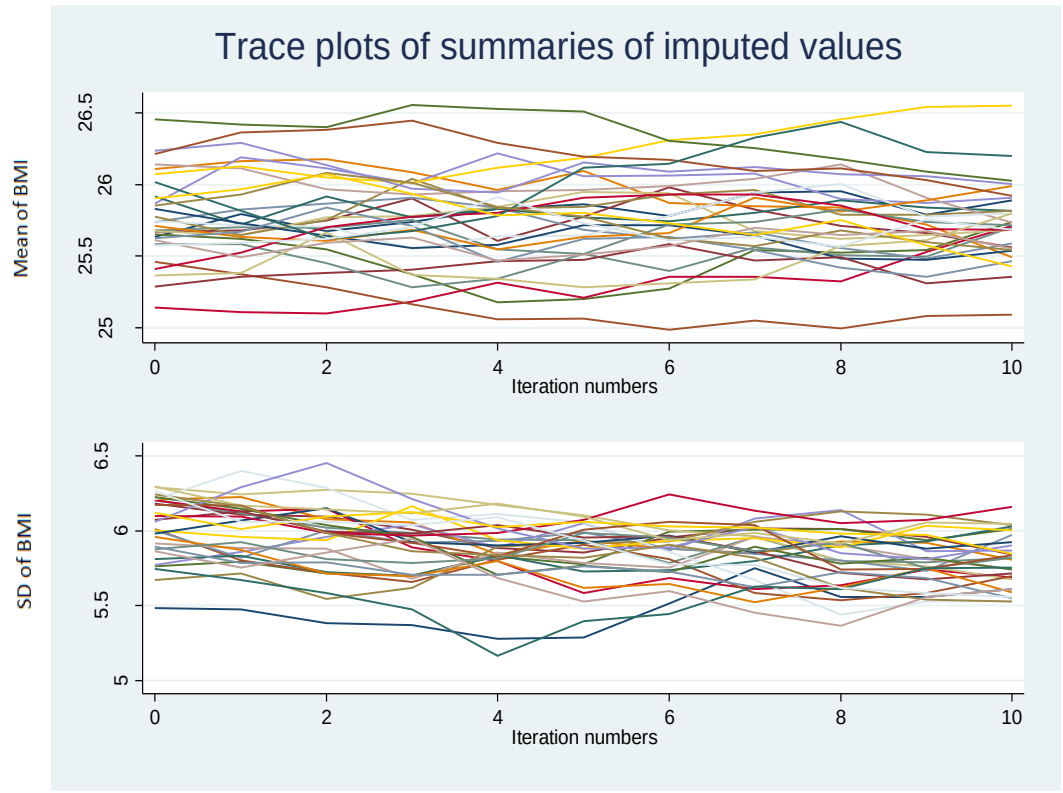
The range for each missing variable was then rechecked, and all were clinically plausible (as shown in Appendix A9.2).

9.3.2.4 Imputation diagnostics

The average RVI of a logistic regression model fitted using the imputed data was 0.0106, which is considered ‘small’, suggesting that missing data had a minimal effect on the variance of the estimate.

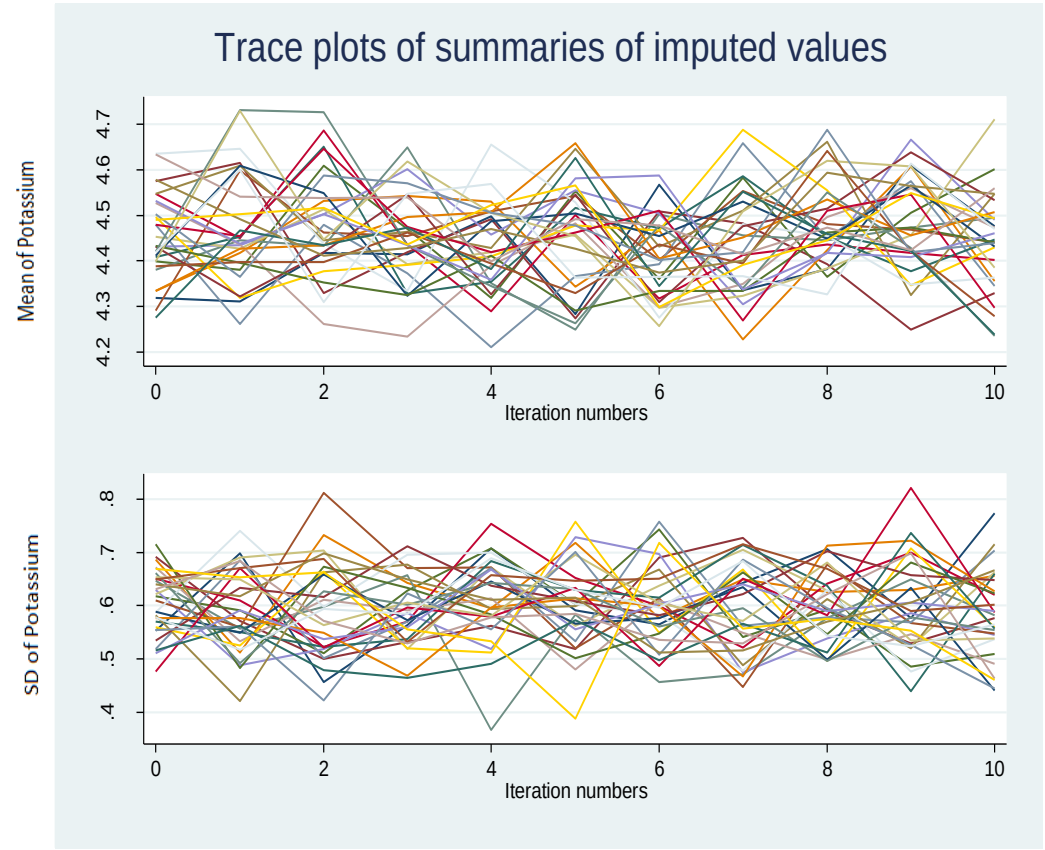
The largest FMI was 0.25, meaning that the number of imputations (30) exceeded the minimum required number (using the rule of thumb that the number of imputations should be at least equal to the highest FMI percentage).

The trace plots for BMI and serum potassium are shown in Figure 11 and Figure 12. There is no apparent trend in any of the iterations, suggesting good convergence of the imputation model.



SD = standard deviation, BMI = body mass index

Figure 11 – Trace plot for body mass index



SD = standard deviation

Figure 12 – Trace plot for serum potassium

9.3.2.5 Multiple imputation sensitivity analysis

The sensitivity analysis involved a comparison of the regression coefficients, standard errors, and *p* values obtained following multivariable logistic regression, using the complete-case and imputed datasets. Initially the analysis was performed using all candidate predictors, and then the models were re-run after omitting BMI (given the large number of missing data for this variable). The results are described below.

Analyses including all candidate predictors

Following the use of 'common sense approaches' to predict missing data (reported in section 9.3.2.1), the number of admissions with one or more missing data point was 379, representing 25.2% of the 1,503 study admissions. The remaining 1,124 admissions were therefore included in the complete-case analyses. The comparison of the multivariable logistic regression estimates for the complete-case and imputed datasets are shown in Table 43.

The differences in the standard errors (between the complete-case and multiply imputed datasets) can be explained by the sample size, with larger standard errors observed for the complete-case analyses. Using a significance level of *p* smaller than 0.157, as selected for MOAT development (discussed in section 9.3.3.1), one predictor that was statistically significant in the complete-case analysis (the use of parenteral medicines) became non-significant following multiple imputation. Conversely, two predictors only became significant following multiple imputation; the use of anti-diabetic medication, and renal function (measured using estimated glomerular filtration rate). As the presumed missingness mechanism was MAR, differences between the two datasets was anticipated. That is to say, only if data were MCAR would the complete cases be a random sub-set of the sample population[60]. While there is no definitive way to distinguish between MAR and MNAR, it has been suggested that observed differences are reviewed to consider if they make sense scientifically[204], with plausible explanations supporting the MAR assumption. I therefore considered whether the change in statistical significance of the three predictors had rational explanations:

- parenteral medicines use – the complete-case analysis suggests that parenteral medicine use reduces the risk of an outcome event (albeit at a nominal significance level), which is counterintuitive given that parenteral administration is known to increase the risk of adverse medication-related outcomes[79 87 91]. The imputed result (i.e. no association) is therefore more plausible;

- anti-diabetic medicine use, and impaired renal function, became statistically significant (in terms of increasing the risk of an outcome event) after multiple imputation. These are logical findings given the recognised risks associated with both[165].

I therefore concluded that it is reasonable to assume that data were MAR.

Table 43 – Comparison of multivariable regression coefficients for complete-case and multiply imputed datasets

Predictor	Complete-cases (observations = 1,124*)			Multiple imputation (observations = 1,503)		
	Regression coefficient†	Standard error	*p* value‡	Regression coefficient†	Standard error	*p* value‡
Demographic						
Age	0.0508	0.048	0.288	0.0397	0.042	0.349
Socioeconomic status	0.0414	0.024	**0.081**	0.0452	0.021	**0.029**
Patient related						
Previous allergy	0.244	0.135	**0.071**	0.261	0.118	**0.027**
Body mass index	-0.00075	0.012	0.950	-0.0030	0.012	0.808
Number of hospital admissions in previous 6 months	0.0586	0.085	0.490	0.0234	0.076	0.757
Primary diagnosis:						
Endocrine and metabolic	Base category		**0.072**	Base category		**0.021**
Nervous system and mental disorders	0.303	0.352		0.319	0.303	
Cardiovascular system	-0.214	0.329		-0.225	0.282	
Respiratory system	-0.460	0.328		-0.316	0.278	
Gastrointestinal system	-0.632	0.369		-0.655	0.316	
Genitourinary system	-0.244	0.348		-0.055	0.301	
Musculoskeletal-intergumentary	-0.0525	0.377		0.041	0.330	
All other categories	-0.0362	0.336		-0.0070	0.286	
Number of comorbidities	0.104	0.037	**0.005**	0.137	0.033	**<0.001**
History of dementia	-0.237	0.223	0.287	-0.264	0.194	0.175
Medicines related						
Number of medicines	0.0211	0.020	0.285	0.0235	0.018	0.189
Use of high-risk medicines:						
Anticoagulants	0.0256	0.172	0.882	0.110	0.154	0.477
Therapeutic heparin	0.452	0.198	**0.022**	0.269	0.179	**0.133**
Anti-diabetic medication	0.131	0.176	0.457	0.222	0.153	**0.146**

Continued from previous page…

Predictor	Complete-cases (observations = 1,124*)			Multiple imputation (observations = 1,503)		
	Regression coefficient†	Standard error	*p* value‡	Regression coefficient†	Standard error	*p* value‡
Opiates	0.0137	0.214	0.949	0.0121	0.197	0.951
Aminoglycosides and glycopeptides	0.434	0.246	**0.078**	0.331	0.226	**0.142**
Other antimicrobials	0.454	0.168	**0.007**	0.362	0.147	**0.014**
Epilepsy medicines	0.316	0.187	**0.092**	0.478	0.165	**0.004**
Antipsychotics	0.195	0.268	0.467	0.164	0.237	0.488
Antiarrhythmics	-0.132	0.223	0.552	-0.0568	0.201	0.777
Antidepressants	0.249	0.153	**0.104**	0.203	0.138	**0.139**
Other high-risk medicines	0.145	0.243	0.549	0.120	0.216	0.580
Parenteral administration	-0.248	0.171	**0.147**	0.0388	0.149	0.795
Laboratory results						
Estimated glomerular filtration rate	-0.0302	0.023	0.188	-0.0366	0.020	**0.069**
Liver disease	-0.0768	0.218	0.725	-0.0995	0.195	0.609
Serum albumin	-0.0069	0.013	0.598	0.00059	0.012	0.959
Serum potassium	-0.0560	0.114	0.622	-0.127	0.101	0.207
Serum sodium	-0.00087	0.013	0.949	-0.0062	0.011	0.581
White cell count	0.0389	0.017	**0.023**	0.0224	0.015	**0.138**
Platelet count	-0.0031	0.008	0.683	0.0028	0.007	0.679

* Following inclusion of values determined using common sense solutions[182]

† Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

‡ Values with a significance level of $p<0.157$ shown in bold

Analyses excluding body mass index

The omission of BMI from the analyses increased the number of observations in the complete-case analysis from 1,124 to 1,440, representing 95.8% of 1,503 admissions. The comparison of the multivariable logistic regression estimates for the complete-case and imputed datasets are shown in Appendix A9.3. Using a significance level of *p* smaller than 0.157, only one predictor (the use of anti-diabetic medication) changed from non-significant to significant. Parenteral administration was non-significant in both datasets, and renal function was significant in both (which are logical findings, as discussed above). As before, given the MAR assumption, some differences between the complete-case and imputed datasets would be expected. It is likely that fewer differences were observed given the higher number of admissions included in the

complete-case analysis. As above, there was no evidence against the MAR assumption.

9.3.3 Model development

Model development is reported in two sections: the selection of predictors during modelling, and model diagnostics.

9.3.3.1 Predictor selection during modelling

Table 44 shows the regression coefficients and *p* values for the candidate predictors following multivariable analysis (using random effects logistic modelling). Results are shown for the full model (containing all candidate predictors), and the BS model.

Backward elimination resulted in the retention of 13 predictors in the model:

- socioeconomic status;
- number of comorbidities;
- number of medicines;
- estimated glomerular filtration rate;
- white cell count;
- previous allergy;
- systemic aminoglycosides and glycopeptides;
- other systemic antimicrobials;
- epilepsy medicines;
- antidepressants;
- primary diagnoses:
    - nervous system and mental disorders;
    - respiratory system;
    - gastrointestinal system.

After considering the sensibility of using the above predictors in the MOAT, I decided to exclude socioeconomic status. The reasons for this were:

1. the relative complexity involved in obtaining socioeconomic status data, requiring:
    - use of the postcode search function provided by the 'English Indices of Deprivation 2015'[149] (which produces an Excel spread sheet containing the deprivation data for the relevant 'neighbourhood');

- extraction of the 'Index of Multiple Deprivation' data from the spread sheet (i.e. the rank position, which ranges from one to 32,844). This is one of 26 separate data fields included in the spread sheet;
- calculation of the neighbourhood's rank position as a percentage of all other English neighbourhoods;

2. the minimal reduction (0.3%) in the model's c-index caused by the removal of socioeconomic status;
3. recognition that inclusion of socioeconomic status may reduce the generalisability of the MOAT (i.e. restrict its use to hospitals in England).

Once socioeconomic status was excluded from the regression model, 'antidepressants' became non-significant, and was therefore also excluded. This left 11 predictors in the BS model (Table 44). I also re-ran the backwards elimination, excluding socioeconomic status from the outset, to assess whether this altered the final predictor selection. The same model was produced.

Table 44 – Multivariable association between predictors and outcome events

Predictor	Full model		Model following backward elimination	
	Regression coefficient† (95% CI)	***p* value‡**	**Regression coefficient† (95% CI)**	***p* value‡**
Demographic				
Age/10 (years)	0.0403 (-0.0446 to 0.125)	0.352	-	-
Socioeconomic status/10, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]*	0.0458 (0.0034 to 0.0882)	0.034	-	-
Patient related				
Previous allergy§	0.266 (0.0193 to 0.512)	0.035	0.318 (0.0691 to 0.566)	0.012
Body mass index (kg/m^2)	-0.0030 (-0.0272 to 0.0213)	0.810	-	-
Number of hospital admissions in previous 6 months	0.0238 (-0.127 to 0.175)	0.757	-	-
Primary diagnosis§:				
Endocrine and metabolic	0.0083 (-0.560 to 0.577)	0.977	-	-
Nervous system and mental disorders	0.331 (-0.129 to 0.791)	0.158	0.414 (0.0183 to 0.810)	0.040
Cardiovascular system	-0.221 (-0.621 to 0.179)	0.279	-	-
Respiratory system	-0.313 (-0.725 to 0.0991)	0.137	-0.274 (-0.577 to 0.0296)	0.077
Gastrointestinal system	-0.656 (-1.177 to -0.136)	0.013	-0.624 (-1.065 to -0.182)	0.006
Genitourinary system	-0.0475 (-0.531 to 0.436)	0.847	-	-
Musculoskeletal-intergumentary systems	0.0491 (-0.494 to 0.592)	0.859	-	-
All other categories	Base category		Base category	
Number of comorbidities	0.139 (0.0646 to 0.214)	<0.001	0.146 (0.0775 to 0.215)	<0.001
History of dementia§	-0.269 (-0.664 to 0.126)	0.182	-	-
Laboratory results				
Estimated glomerular filtration rate/10 (ml/min/1.73m^2)	-0.0372 (-0.0781 to 0.0038)	0.076	-0.0360 (-0.0734 to 0.0014)	0.059
Liver disease§	-0.101 (-0.490 to 0.287)	0.609	-	-
Serum albumin (g/L)	0.00063 (-0.0225 to 0.0237)	0.958	-	-
Serum potassium (mmol/L)	-0.129 (-0.332 to 0.0742)	0.213	-	-
Serum sodium (mmol/L)	-0.0062 (-0.0286 to 0.0161)	0.583	-	-
White cell count (10^9/L)	0.0227 (-0.0077 to 0.0531)	0.143	0.0274 (-0.0008 to 0.0557)	0.057
Platelet count/10 (10^9/L)	0.0028 (-0.0105 to 0.0161)	0.677	-	-

Continued from previous page…

Predictor	Full model		Model following backward elimination	
	Regression coefficient† (95% CI)	***p* value‡**	**Regression coefficient† (95% CI)**	***p* value‡**
Medicines related				
Number of medicines	0.0239 (-0.0122 to 0.0599)	0.194	0.0406 (0.0074 to 0.0737)	0.016
Use of high-risk medicines§				
Anticoagulants	0.111 (-0.197 to 0.419)	0.479	-	-
Therapeutic heparin	0.273 (-0.0888 to 0.635)	0.139	-	-
Anti-diabetic medication	0.224 (-0.0816 to 0.530)	0.151	-	-
Opiates	0.0122 (-0.379 to 0.403)	0.951	-	-
Aminoglycosides and glycopeptides	0.335 (-0.118 to 0.787)	0.147	0.387 (-0.0535 to 0.828)	0.085
Other antimicrobials	0.3670 0.0641 to 0.670)	0.018	0.364 (0.0909 to 0.637)	0.009
Epilepsy medicines	0.486 (0.135 to 0.837)	0.007	0.450 (0.1111 to 0.789)	0.009
Antipsychotics	0.165 (-0.305 to 0.635)	0.492	-	-
Antiarrhythmics	-0.0583 (-0.459 to 0.342)	0.775	-	-
Antidepressants	0.207 (-0.0737 to 0.488)	0.148	-	-
Other high-risk medicines	0.120 (-0.310 to 0.550)	0.585	-	-
Parenteral administration§	0.0406 (-0.258 to 0.339)	0.790	-	-
Constant	-0.668		-1.674	

* Deprivation rank based on patients' postcode, shown as the ranked position as a percentage of all neighbourhoods in England (where 1 is the most deprived)

† Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

‡ Test for difference between admissions with and without occurrence of outcome event. Obtained from multivariable regression modelling

§ Categorical exposure variable. For the purposes of calculating the predicted risk for individual patients (as described in section 9.2.6.2), categorical variables were coded as 'one' if present and 'zero' if absent

Age, socioeconomic status, estimated glomerular filtration rate, and platelet count analysed as deciles to aid interpretability (i.e. after dividing the actual value by ten).

To permit direct comparison between the univariable and multivariable results (shown in section 9.3.1.5) the odds ratio for each predictor in the full model was also calculated (Appendix A9.4).

9.3.3.2 Model diagnostics

The results for the model diagnostics are presented in four sections: evidence for clustering caused by duplicate admissions, the accuracy of the 'quadrature approximation' used to estimate the random effects model, the check for specification error, and detection of outlying observations. Each is described below.

Evidence of clustering

The estimates of the correlation within clusters (rho) produced after fitting the BS model in both the non-imputed and multiply imputed datasets is shown in Table 45. The result of the likelihood ratio test (performed following the complete-case analysis of non-imputed data) is also shown.

Table 45 – Evidence of clustering following random effects modelling using predictors selected by backward elimination

	Random effects model (predictors selected by backward elimination)	
	Non-imputed dataset	**Multiply imputed dataset**
Number of observations (i.e. number of admissions included in the analysis)	1,494	1,503
Number of groups (i.e. number of individual patients included in the analysis)	1,436	1,444
Rho	0.000049	0.030
Likelihood ratio test of rho = 0	p = 0.498	N/A

Rho = correlation within clusters, N/A = not applicable

This shows that 59 of the 1,503 observations (admissions) included in the multiply imputed model were patients included in the study more than once (i.e. the number of admission minus the number of groups). For the complete case analysis (using non-imputed data), 1,494 observations were included, of which 58 patients were admitted more than once.

Rho was close to zero following the complete-case analysis, with the likelihood ratio test suggesting no evidence for clustering (p = 0.498). The value of rho was larger for the multiply imputed model (0.03), which may be due to the slightly higher number of observations included in this model; this suggests some evidence of clustering, albeit relatively small[205].

In summary, there is some evidence that duplicate admissions caused within patient correlation, providing support for the use of random effects modelling.

Accuracy of the quadrature approximation

Following the use of 'quadchk' to check the adequacy of the quadrature approximation of the BS model, the largest relative difference in the regression coefficients of the predictors was 0.0002%. This indicates that the results may be confidently interpreted[188].

Testing for specification error

The Stata output for the specification error test is shown in Table 46. This shows that the 'predicted value' (_hat) was statistically significant, and the 'prediction squared' (_hatsq) was not significant. This suggests there is no evidence for a specification error in the BS model.

Table 46 – Stata output for specification error check of backward selection model

.

study_outcome	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_hat	.9010783	.1057824	8.52	0.000	.6937487	1.108408
_hatsq	-.1126027	.0958771	-1.17	0.240	-.3005184	.075313
_cons	.0342448	.0692424	0.49	0.621	-.1014678	.1699575

_hat = 'predicted value', _hatsq = 'predicted value squared'; _cons = constant, Coef = regression coefficient, Std Err = standard error, t and P>|t| provide the Student's *t* and 2-tailed *p* value

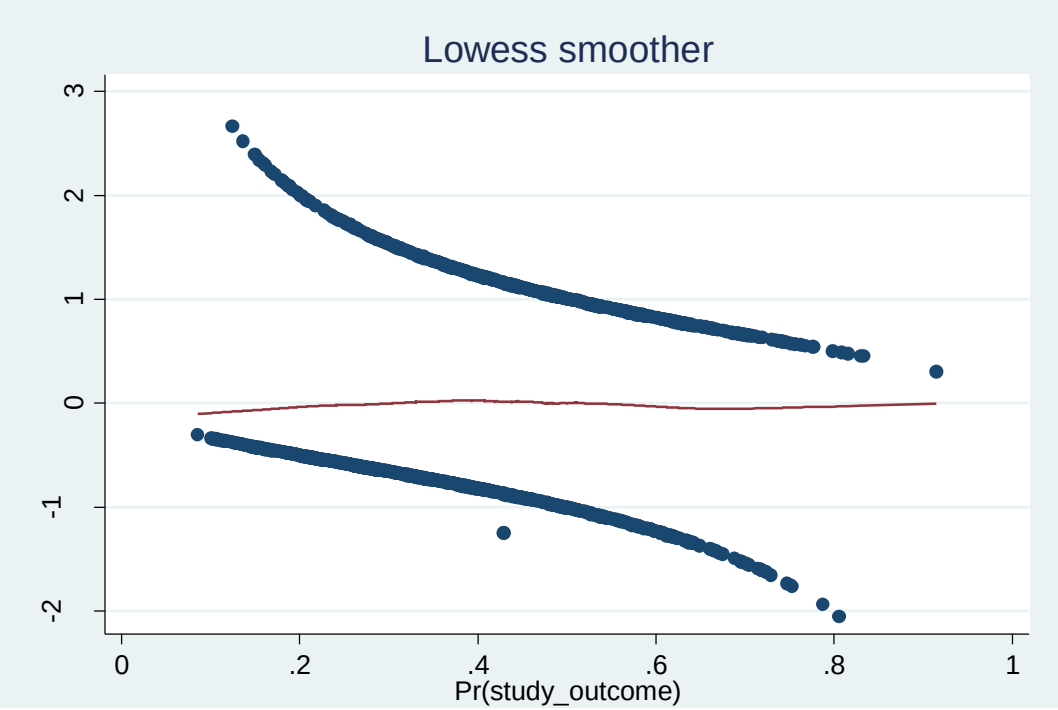
Detection and review of outlying observations

Plots for the residual, influence measure, and diagnostic statistics are presented below, followed by a summary of the impact of the outlying observations (cases).

Standardised Pearson and deviance residuals

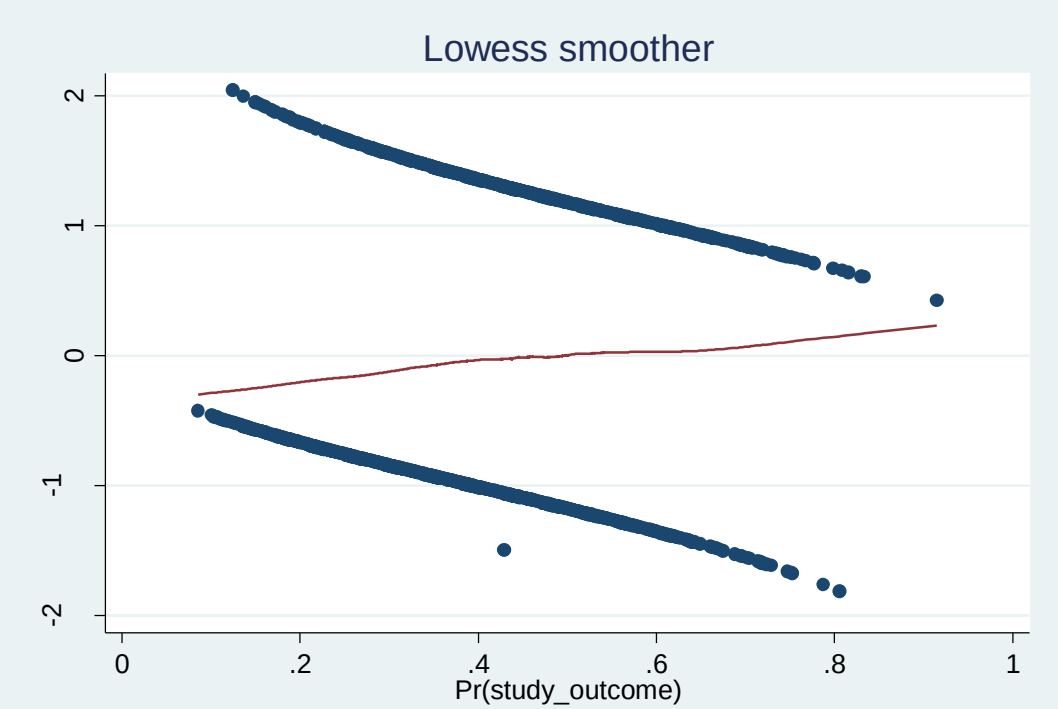
Plots of the standardised Pearson and deviance residuals against the predicted risk probabilities are shown in Figure 13 and Figure 14 respectively. Eighteen cases had a standardised Pearson residual of above two (ranging from 2.06 to 2.66), and one case had a standardised Pearson residual below minus two (-2.06). Only one case had a deviance residual outside the range of two to minus two (2.05).

A lowess line was used to create smooth lines through the scatterplots. Where a model is correct, with no significant incorporation of potential outliers, a lowess smooth of the plot of the residuals against the logistic probability should result in approximately a horizontal line with zero intercept; significant departure from this suggests that potential outliers may have dramatic impact on the fit of the model[189]. No significant departure was observed for either the standardised Pearson or deviance residuals.



Pr(study-outcome) = predicted probability of study outcome, lowess = locally weighted scatterplot smoothing

Figure 13 – Standardised Pearson residual against estimated logistic probability of study outcome

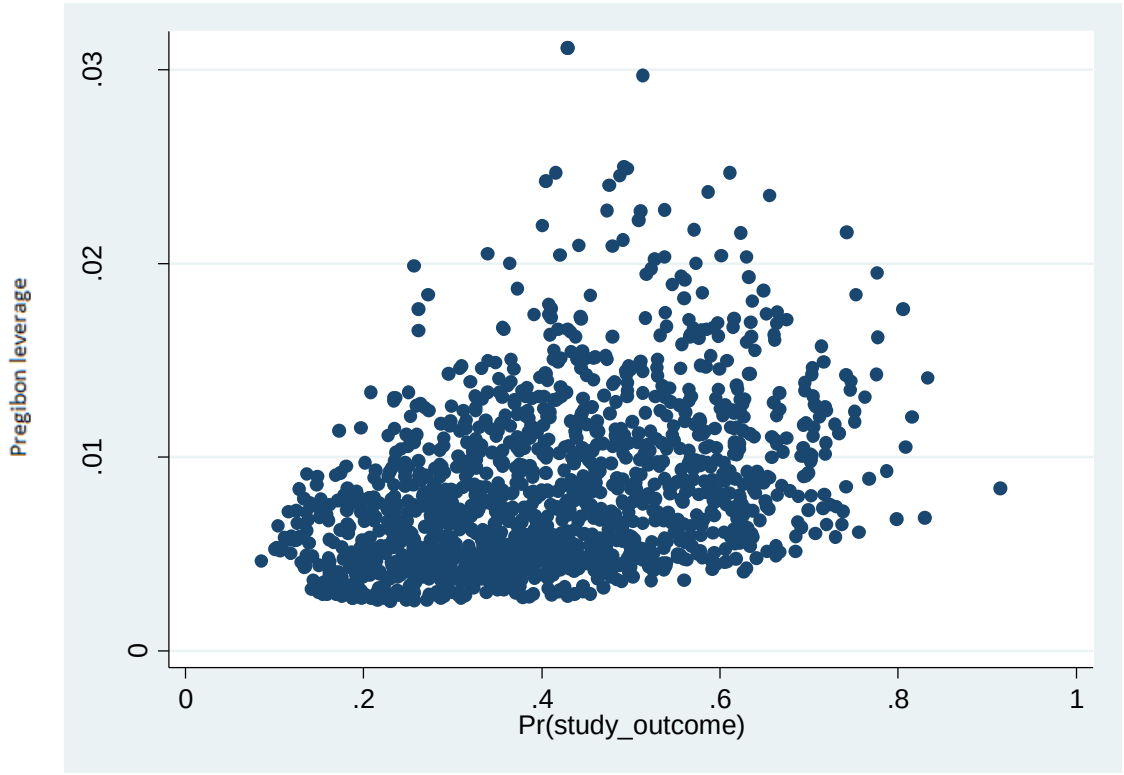


Pr(study-outcome) = predicted probability of study outcome, lowess = locally weighted scatterplot smoothing

Figure 14 – Deviance residual against estimated logistic probability of study outcome

Pregibon leverage

The plot of Pregibon leverage against the predicted risk probabilities is shown in Figure 15. The average value for Pregibon leverage was 0.008, therefore a cut-off of 0.024 was used to identify outlying cases. Nine outlying cases were found (ranging from 0.0243 to 0.031)

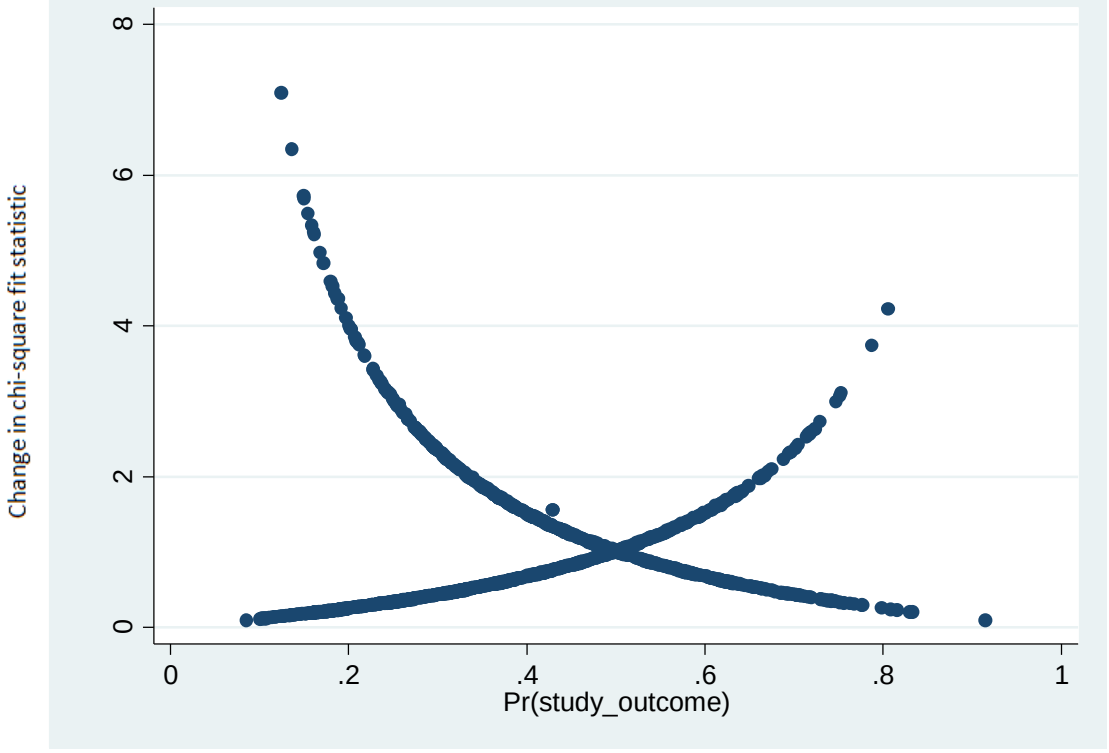


Pr(study-outcome) = predicted probability of study outcome

Figure 15 – Pregibon leverage against estimated logistic probability of study outcome

Change in chi-square fit statistic

The plot of the change in chi-square fit statistic against the predicted risk probabilities is shown in Figure 16. Twenty four outlying cases were found (ranging from 3.847 to 7.094)

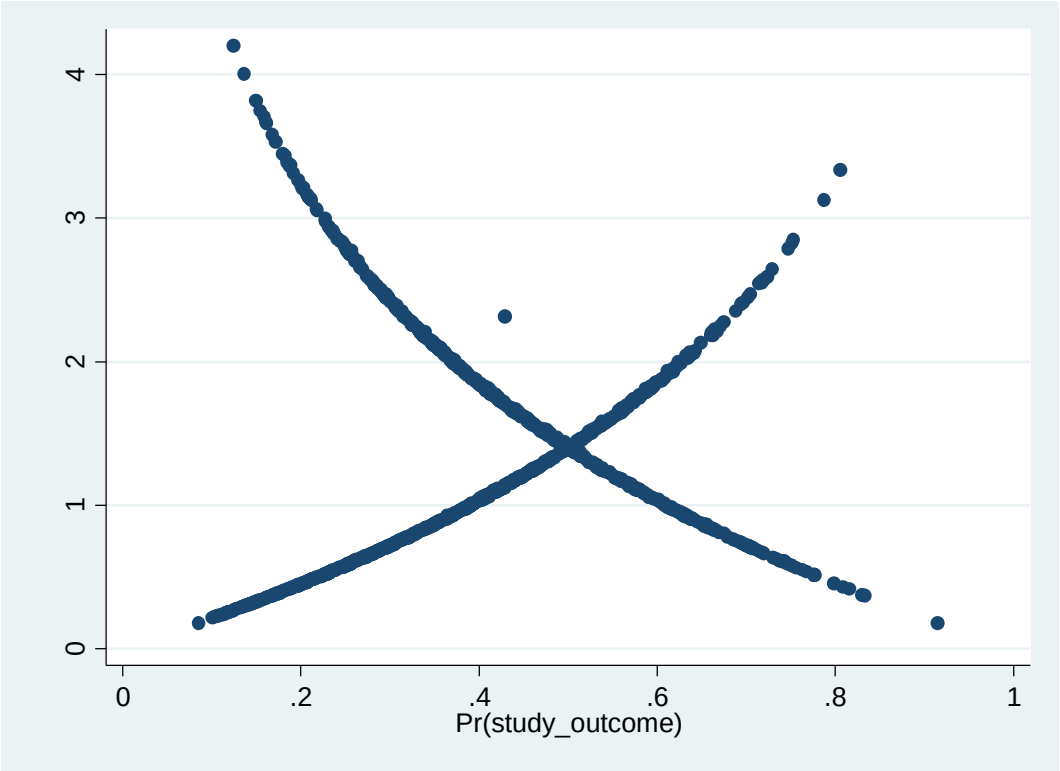


Pr(study-outcome) = predicted probability of study outcome

Figure 16 – Change in chi-square fit statistic against estimated logistic probability of study outcome

Change in deviance statistic

The plot of the change in deviance statistic against the predicted risk probabilities is shown in Figure 17. Two outlying cases were found (4.00 and 4.20)

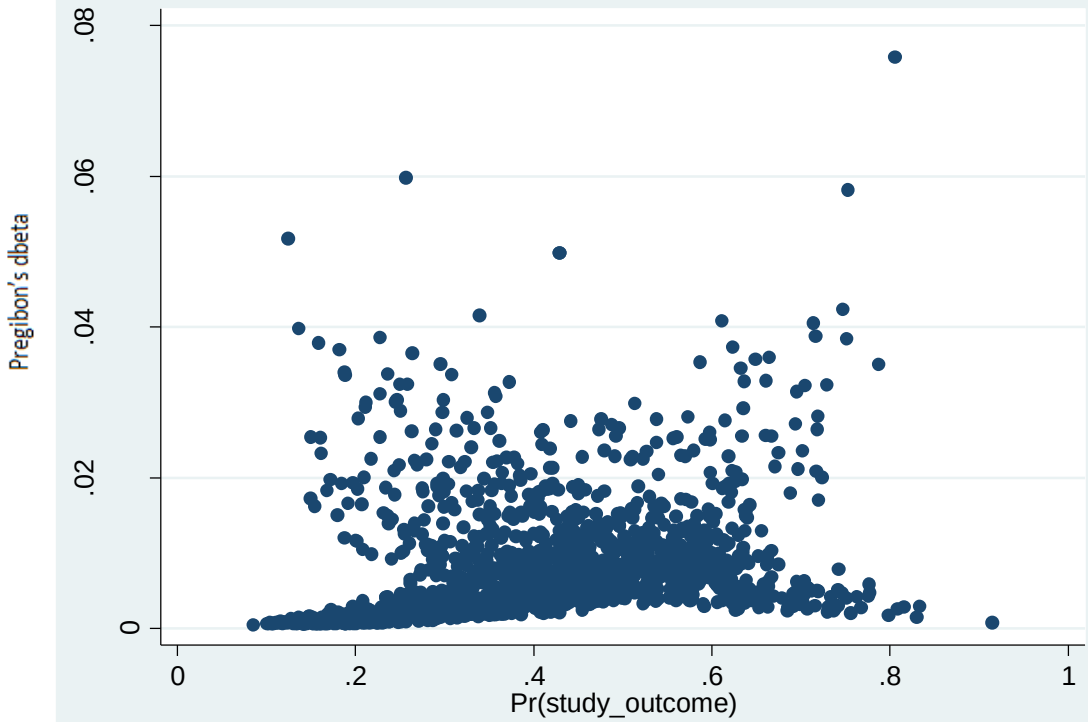


Pr(study-outcome) = predicted probability of study outcome

Figure 17 – Change in deviance statistic against estimated logistic probability of study outcome

Pregibon's dbeta

The plot of the change in regression coefficients (Pregibon's dbeta) against the predicted risk probabilities is shown in Figure 18. Using a cut-off of 0.04 (selected following visual examination of the data), 10 cases were found to be outliers (ranging from 0.0405 to 0.0758)



Pr(study-outcome) = predicted probability of study outcome

Figure 18 – Pregibon's dbeta against estimated logistic probability of study outcome

Summary

The review of the residuals, influence measure, and diagnostic statistics identified 38 outlying cases (Table 47), Outlying values are shown in bold.

Table 47 – Summary of residual, influence measure and diagnostic statistics for the backward selection model

Study identifier	Residuals		Pregibon leverage	Change in chi-square fit statistic	Change in deviance statistic	Pregibon's dbeta
	Standardised Pearson	Deviance				
4	**2.06***	1.82	0.0039	**4.23**	3.32	0.0166
32	-1.27	-1.37	**0.0247**	1.61	1.94	**0.0408**
193	1.02	1.18	**0.0249**	1.04	1.44	0.0266
207	1.99	1.79	0.0046	**3.96**	3.21	0.0185
270	**2.09**	1.83	0.0076	**4.36**	3.37	0.0336
312	**2.20**	1.88	0.0041	**4.83**	3.53	0.0197
469	**2.13**	1.85	0.0081	**4.53**	3.43	0.0369
526	**2.11**	1.84	0.0043	**4.43**	3.39	0.0192
528	-1.73	-1.66	0.0139	2.99	2.79	**0.0423**
542	-0.83	-1.02	**0.0243**	0.70	1.06	0.0173
552	**2.23**	1.89	0.0037	**4.97**	3.58	0.0182
649	1.99	1.79	0.007	**3.95**	3.21	0.0278
654	1.72	1.65	0.0199	2.95	2.77	**0.0598**
661	-1.59	-1.58	0.0157	2.54	2.54	**0.0405**
667	1.41	1.47	0.0205	1.99	2.21	**0.0415**
672	1.04	1.20	**0.0245**	1.08	1.47	0.0270
693	-1.00	-1.16	**0.0250**	0.99	1.39	0.0255
744	1.96	1.77	0.0043	**3.85**	3.16	0.0165
835	**2.28**	1.91	0.0044	**5.22**	3.66	0.0232
865	**2.31**	1.92	0.007	**5.33**	3.71	0.0378
884	-1.25	-1.50	**0.0311**	1.55	2.31	**0.0498**
900	**2.09**	1.83	0.0077	**4.36**	3.37	0.0340
964	-1.25	-1.50	**0.0311**	1.55	2.31	**0.0498**
1004	**2.09**	1.83	0.0028	**4.35**	3.36	0.0120
1078	0.99	1.15	**0.0297**	0.97	1.37	0.0298
1091	-0.85	-1.04	**0.0247**	0.73	1.10	0.0184
1097	**2.03**	1.80	0.0047	**4.10**	3.27	0.0192
1100	**2.66**	**2.04**	0.0072	**7.09**	**4.20**	**0.0518**
1121	**2.39**	1.95	0.0044	**5.70**	3.81	0.0254
1179	**-2.05**	-1.81	0.0176	**4.22**	3.34	**0.0758**
1232	-1.76	-1.67	0.0184	3.11	2.85	**0.0582**
1277	2.00	1.79	0.0029	**4.00**	3.22	0.0116
1281	1.99	1.79	0.0029	**3.96**	3.21	0.0115
1296	**2.34**	1.93	0.003	**5.49**	3.75	0.0163
1323	**2.39**	1.95	0.003	**5.72**	3.82	0.0173
1341	**2.52**	1.99	0.0062	**6.35**	**4.00**	0.0398
1406	**2.14**	1.85	0.0033	**4.59**	3.45	0.015
1410	**2.29**	1.91	0.0048	**5.25**	3.67	0.0253

* Outlying values shown in bold

To assess the impact of excluding the outlying cases from the regression analysis I compared the regression coefficients and *p* values from a model fitted using all cases, with a model where all outlying cases were excluded (Table 48). Removal of the 38 cases caused the regression coefficients to change by an average of 24.7% (range 1.6% to 42.3%), but did not change the overall significance of the predictors (i.e. all remained 'significant' using a significance level of *p* smaller than 0.157, as selected for MOAT development). Removal also caused the model chi-square to increase from 149.9 to 205.4, indicating an increase in predictive information.

Table 48 – Comparison of parameter estimates for regression models with and without outlying cases

Predictor	Model including all cases		Model excluding outlying cases		Percentage change in regression coefficients‡
	Regression coefficient*	*p* value†	Regression coefficient*	*p* value†	
Number of comorbidities	0.14	<0.001	0.172	<0.001	22.9
Number of medicines	0.041	0.011	0.055	0.001	34.1
Estimated glomerular filtration rate/10 (ml/min/1.73m^2)	-0.04	0.029	-0.056	0.003	40.0
White cell count (10^9/L)	0.027	0.053	0.038	0.008	40.7
Previous allergy	0.297	0.010	0.344	0.004	15.8
Nervous system and mental disorders	0.395	0.040	0.562	0.005	42.3
Respiratory system	-0.284	0.056	-0.327	0.034	15.1
Gastrointestinal system	-0.577	0.006	-0.586	0.011	1.6
Anticoagulants	0.404	0.063	0.483	0.037	19.6
Other antimicrobials	0.345	0.008	0.425	0.002	23.2
Epilepsy medicines	0.448	0.005	0.521	0.002	16.3

* Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

† Test for difference between admissions with and without occurrence of outcome event. Obtained from multivariable regression modelling

‡ Comparison between models fitted with and without the outlying cases

Deleting cases with large residuals or more extreme values almost always improves the fit of model[189], but is not appropriate if the observations are valid. To establish possible explanations (for the case being an outlier), I therefore reviewed the predicted risk probabilities of the outlying cases, and whether each of these cases experienced an outcome event (Table 49).

Table 49 – Predicted risk probabilities and occurrence of the outcome event for outlying cases

Study identifier of outlying case	Predicted risk probability	Occurrence of outcome event
4	0.19	Yes
32	0.61	No
193	0.50	Yes
207	0.20	Yes
270	0.19	Yes
312	0.17	Yes
469	0.18	Yes
526	0.19	Yes
528	0.75	No
542	0.41	No
552	0.17	Yes
649	0.20	Yes
654	0.26	Yes
661	0.71	No
667	0.34	Yes
672	0.49	Yes
693	0.49	No
744	0.21	Yes
835	0.16	Yes
865	0.16	Yes
884	0.43	No
900	0.19	Yes
964	0.43	No
1004	0.19	Yes
1078	0.51	Yes
1091	0.42	No
1097	0.20	Yes
1100	0.12	Yes
1121	0.15	Yes
1179	0.81	No
1232	0.75	No
1277	0.20	Yes
1281	0.20	Yes
1296	0.16	Yes
1323	0.15	Yes
1341	0.14	Yes
1406	0.18	Yes
1410	0.16	Yes

This shows that 28 of the 38 outliers experienced an outcome event, the majority of whom had low predicted risk probabilities. For example case ‘1100’ had no comorbidities, was prescribed only two medicines, had ‘normal’ renal function, a white

cell count within the standard reference range, and was not prescribed any of the three high-risk medicines included in the BS model. Despite his low predicted risk probability (of 0.12), this patient experienced a valid outcome event; they were prescribed a sub-therapeutic dose of prophylactic heparin for thromboprophylaxis (based on their weight), resulting in an increased risk of thromboembolism.

The remaining ten cases did not experience an outcome event despite having relatively high predicted risk probabilities. Given the potential under-reporting of MRPs by study pharmacists (as discussed in chapter 7), it was possible that these patients may have experienced an outcome event that was not reported. I therefore re-ran the regression excluding the outlying cases that were recorded as not having experienced an outcome event. This resulted in a small average change in the regression coefficients (of 3.8%), and a modest increase in the model chi-square compared to the model with all cases (from 149.9 to 159.7).

After considered the evidence presented above, I chose not to exclude the outlying cases from the analysis for the following reasons:

- although the outlying cases had some impact on the regression coefficients and model fit, this was not sufficient to impact on the statistical significance of any of the model predictors (i.e. all predictors would be retained in the model irrespective of whether outliers included);
- the cases that experienced an outcome event despite having low predicted risk probabilities appear to be genuine findings;
- exclusion of patients with high predicted risk probabilities who did not experience a study event resulted in only minor changes in the model's parameters.

An additional consideration was the fact that while exclusion of the outlying cases would result in increased model fit, it may cause overfitting, with overestimation of the predictive performance of the model when used in a new set of patients.

9.3.4 Assessing model performance

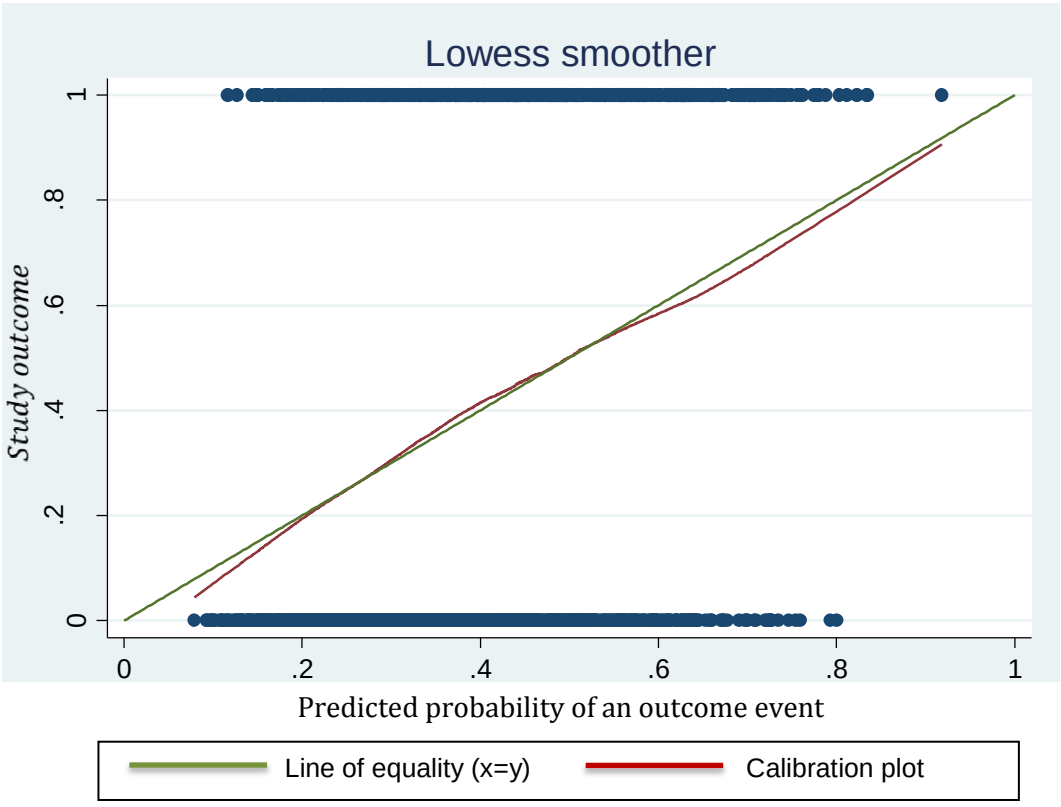
The discrimination and calibration of the BS model are reported below. The optimism-corrected discrimination is reported in section 9.3.5.

Discrimination

The c-index for the BS model was 0.681 (95% confidence interval 0.654 to 0.708).

Calibration

The calibration plot for the BS model is shown in Figure 19. This is a plot of the predicted probability of an outcome event against observed outcome frequency. The smoothed (lowess) line, and line of equality (i.e. the 45 degree line) are shown. This plot suggests slight overestimation of the risk of an outcome event for patients at lower risk (i.e. predicted risk probability under 20%), and those at higher risk (over 60%), but the calibration slope suggests good overall calibration (slope 0.974, intercept 0.012).



Study outcome = observed outcome frequency

Figure 19 – Calibration plot of predicted probability of an outcome event against the proportion of admissions that experienced an event

9.3.5 Internal validation / adjustment for optimism

Discrimination

Following bootstrap validation the estimated optimism for the c-index was found to be 0.027 (range -0.021 to 0.0632, standard deviation 0.014). Subtracting this value from the corresponding BS model, that is to say, prior to the removal of 'non-sensible' predictors (c-index 0.684), the estimated optimism-corrected c-index for the BS model was 0.657.

Calibration

Bootstrap validation suggested slight overfitting of the BS model, with an average calibration slope of 0.855 (range 0.703 to 1.126, standard deviation 0.066). The optimism for the calibration slope was therefore used as a linear shrinkage factor to adjust the regression coefficients of the BS model (Table 50).

Table 50 – Multivariable association between predictors and outcome events (backward elimination model) before and after correction for optimism

Predictor	Regression coefficient* (backward elimination model)	Regression coefficient* (following correction for optimism†)
Number of comorbidities	0.146	0.125
Estimated glomerular filtration rate/10 (ml/min/1.73m^2)	-0.0360	-0.0308
White cell count (10^9/L)	0.0274	0.0234
Number of medicines	0.0406	0.0347
Previous allergy§	0.318	0.272
Nervous system and mental disorders§	0.414	0.354
Respiratory system§	-0.274	-0.234
Gastrointestinal system§	-0.624	-0.533
Aminoglycosides and glycopeptides§	0.387	0.331
Other antimicrobials§	0.364	0.311
Epilepsy medicines§	0.450	0.385
Constant	-1.674	

* Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

† Original regression coefficients corrected by uniform linear shrinkage factor (0.855)

§ Categorical exposure variable. For the purposes of calculating the predicted risk for individual patients (as described in section 9.2.6.2), categorical variables were coded as 'one' if present and 'zero' if absent

9.3.6 Development of a decision aid (the MOAT)

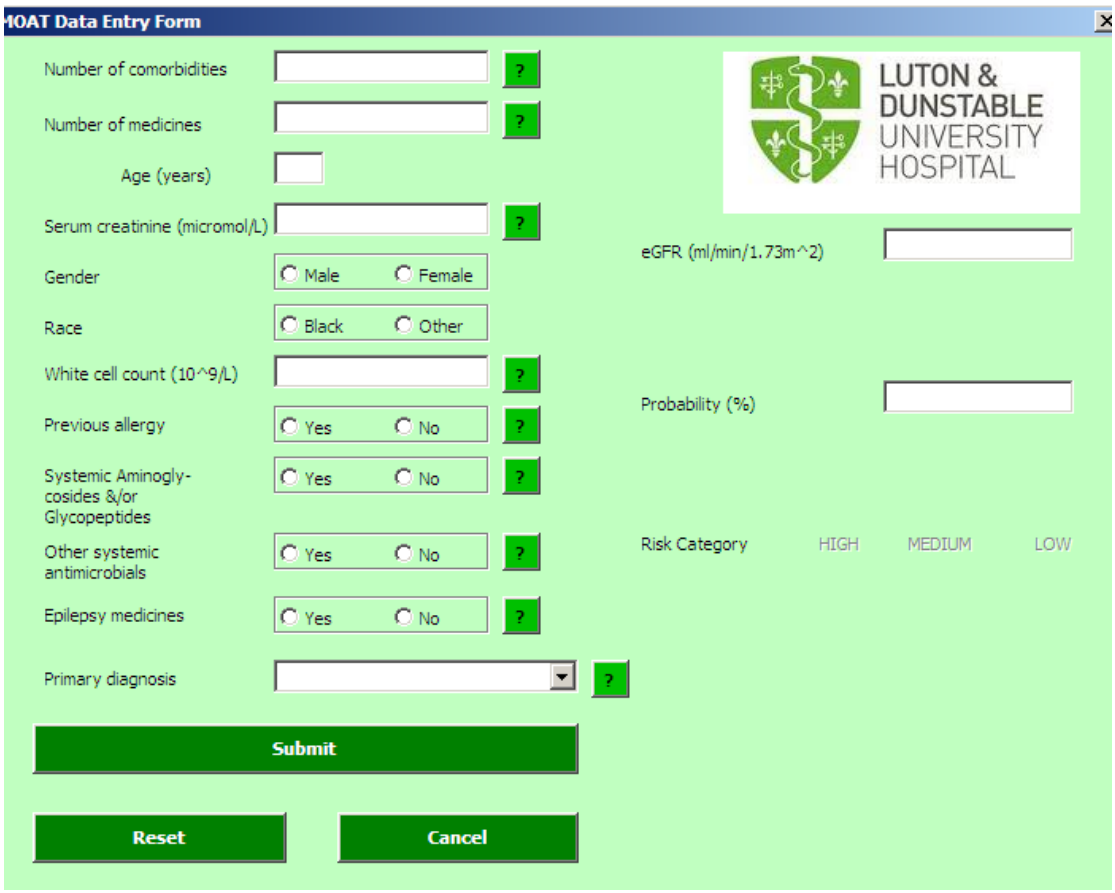
The development of the MOAT is reported in three sections: the presentation format, creation of risk groups, and assessment of clinical usefulness.

9.3.6.1 Presentation format

The MOAT was developed as a Microsoft Excel data entry sheet (Figure 20) that calculates individual patients' probability of an outcome event. This requires pharmacy staff to input data for all candidate predictors, including data required to estimate renal function using the modified Modification of Diet in Renal Disease (MDRD) equation[121]:

- number of comorbidities;
- number of 'regular' medicines prescribed to be given on the first full day of admission to hospital;
- white cell count, 10^9/L (first documented result following admission);
- Previous allergy (yes or no);
- systemic aminoglycosides and glycopeptides use (yes or no);
- use of systemic antimicrobials other than aminoglycosides and glycopeptides (yes or no);
- use of epilepsy medicines (yes or no);
- primary diagnosis of 'nervous system and mental disorders' (yes or no);
- primary diagnosis of 'respiratory system' (yes or no);
- primary diagnosis of 'gastrointestinal system' (yes or no);
- data required to estimate renal function:
    - serum creatinine, micromoles per litre (first documented result following admission);
    - age in years (at admission to hospital);
    - gender (male or female);
    - ethnicity (black or non-black).

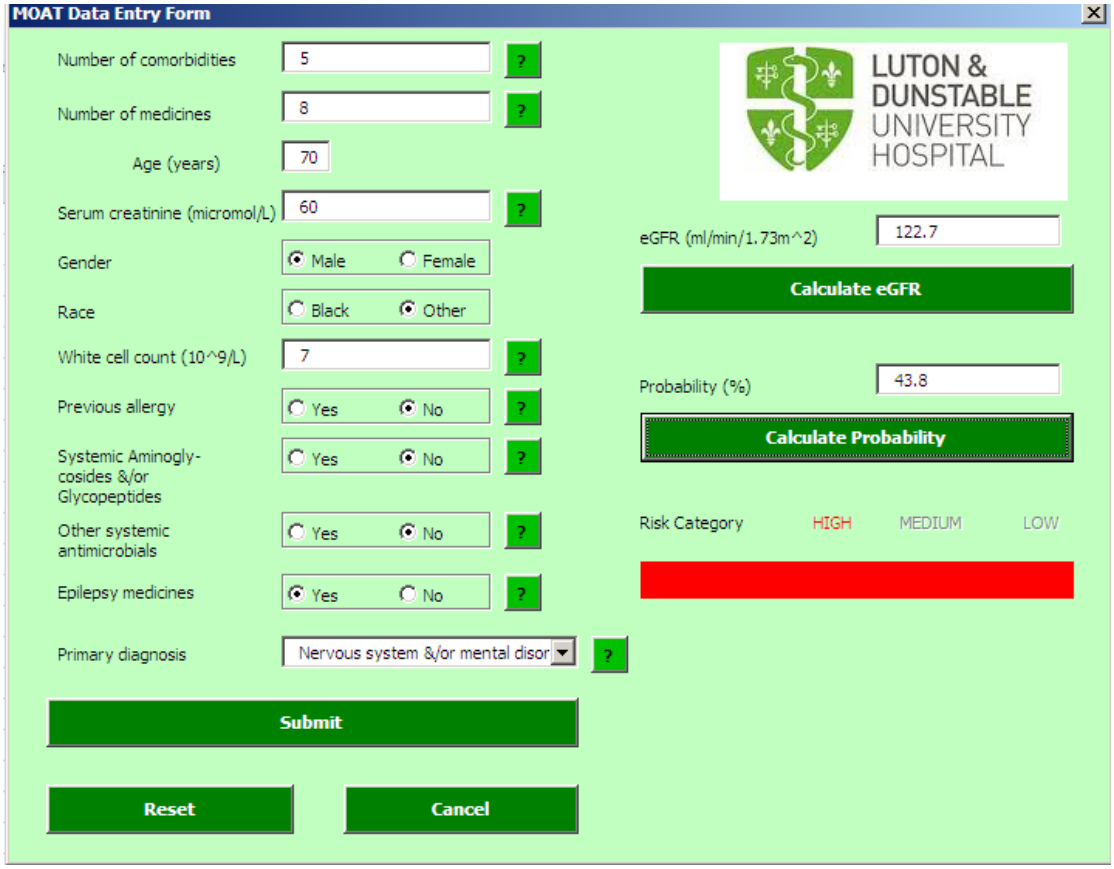
Additional guidance is provided using 'help links'.



eGFR = estimated glomerular filtration rate

Figure 20 – Screenshot of Medicines Optimisation Assessment Tool (MOAT) data entry sheet (prior to data entry)

The MOAT then calculates the estimated glomerular filtration rate (eGFR), and the predicted probability of experiencing an MSP MRP. The probability is shown as the percentage probability, in addition to the patient's risk category (high, medium or low-risk, which is colour coded as red, amber or green respectively). An example of a completed MOAT assessment is shown in Figure 21.



eGFR = estimated glomerular filtration rate

Figure 21 – Screenshot of Medicines Optimisation Assessment Tool (MOAT) data entry sheet (following data entry)

9.3.6.2 Creation of risk groups

Three risk categories were created, high, medium and low.

The decision threshold to distinguish between low and medium-risk patients was informed by the target sensitivity for the MOAT. As discussed in section 9.2.1, an acceptable target sensitivity was established by including a question in the survey of healthcare professionals and patient / public representatives (discussed in chapter 5); as a result, a target sensitivity of 90% was chosen. The decision threshold was therefore selected as the predicted risk probability that corresponded to a sensitivity of 90%; this was approximately 0.25 (i.e. 25%). Patients below this cut-off were categorised as low-risk, and patients above this cut-off were categorised as medium-risk (Table 51). Using this decision threshold, the sensitivity of the MOAT was 89.9% (95% confidence interval 87.6% to 92.4%), and specificity 30.2% (95% confidence interval 27.2% to 33.2%).

Table 51 – MOAT outcomes using the decision threshold between low and medium-risk categories

		Occurrence of moderate or severe preventable medication related problem		Total (n)
		Yes (n)	No (n)	
MOAT outcome	Screen positive (n)	545	620	1165
	Screen negative (n)	61	268	329
Total (n)		606	888	1494*

* Excludes nine admissions with missing data

A decision threshold to distinguish between medium and high-risk patients was informed by considering workload pressures. Following discussion with a group of practising pharmacy staff, it was decided to choose a threshold equivalent to pharmacy staff needing to see 50% of patients; this corresponded to a predicted risk probability of approximately 0.35 (35%). As shown in Table 52, based on this decision threshold, the MOAT's sensitivity was 66.2% (95% confidence interval 62.4% to 70.0%), and specificity 61.0% (95% confidence interval 57.8% to 64.2%).

Table 52 – MOAT outcomes using the decision threshold between medium and high-risk categories

		Occurrence of moderate or severe preventable medication related problem		Total (n)
		Yes (n)	No (n)	
MOAT outcome	Screen positive (n)	401	346	747
	Screen negative (n)	205	542	747
Total (n)		606	888	1494*

* Excludes nine admissions with missing data

The predicted probability range for the risk groups are therefore:

- low – lower than 25.3% risk of experiencing an outcome event;
- medium – between 25.3 and 34.9% risk of experiencing an outcome event;
- high – greater than or equal to 35.0% risk of experiencing an outcome event.

9.3.6.3 Assessment of clinical usefulness

The decision curve for the MOAT is shown in Figure 22. As anticipated, the 'treat none' and 'treat all' lines cross at the prevalence of the outcome event (40.6%). The MOAT is comparable to the strategy of 'treat all' at low threshold probabilities, and comparable

to ‘treat none’ at high probabilities. This is because the probability of an outcome event predicted by the MOAT ranges from 0.09 (9%) to a maximum of 0.86 (86%); using the MOAT below or above this range therefore gives the same result as ‘treat all’ or ‘treat none’. Between approximately 0.70 (70%) and 0.85 (85%) the net benefit is approximately equal to the strategy of ‘treat none’, this is because of the relative increase in false positive compared to true positive results. Between approximately 0.15 (15%) and 0.70 (70%) the MOAT is better than both the ‘treat none’ and ‘treat all’ strategies, suggesting it is of value for patients with predicted risk probabilities within this range[200].

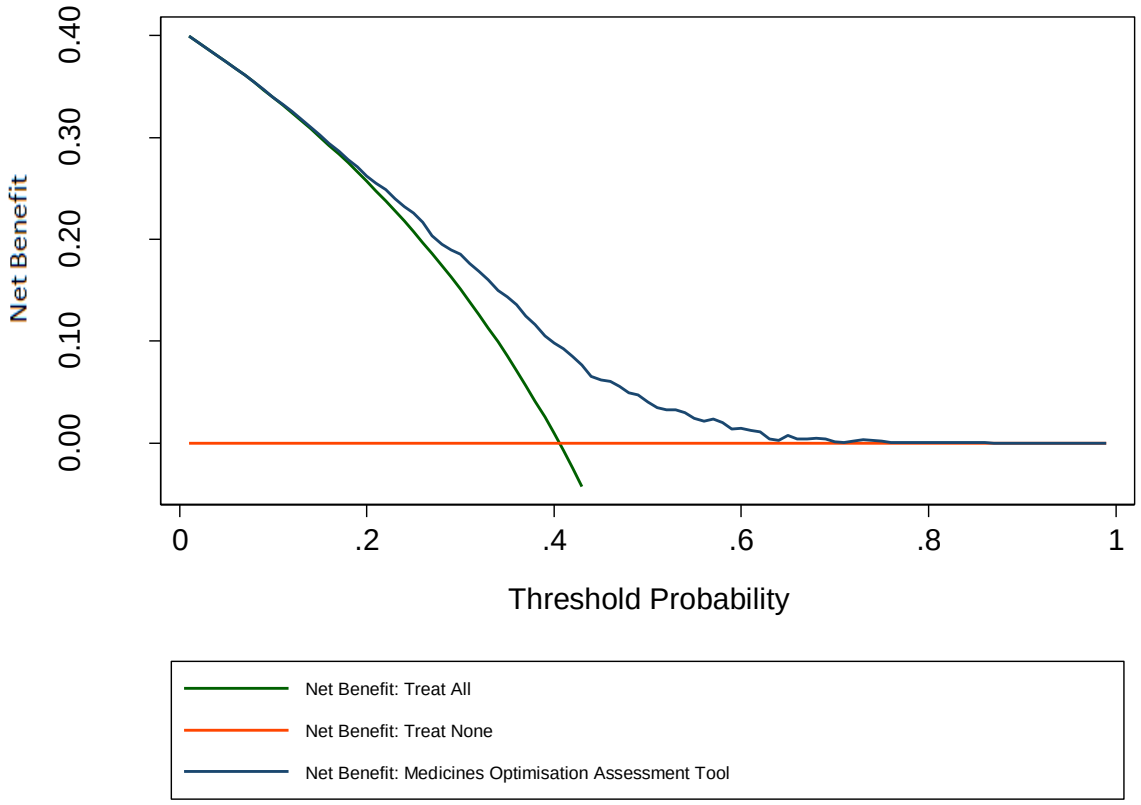


Figure 22 – Decision curve for the Medicines Optimisation Assessment Tool (MOAT)

Given that the decision thresholds selected for the MOAT (section 9.3.6.2) were 0.25 (25%) and 0.35 (35%), both are within the range of probability where the MOAT is considered to be clinically useful. Should a higher decision threshold be selected (due to extreme work pressures), the MOAT would continue to be of value in terms of clinical decision making, up to a predicted risk probability of approximately 0.7 (70%), suggesting significant flexibility.

9.4 Discussion

Key findings

Multivariable logistic regression modelling was used to establish the relationship between pre-selected candidate predictors and patients with an outcome event, namely at least one MSP MRP. Backwards elimination was then used to produce a parsimonious model, aiming to increase clinical applicability while retaining reasonable predictive performance. The model was then adjusted for optimism (to reduce the potential for overconfident predictions when applied to a new group of patients), and the adjusted regression equation used to develop an electronic decision aid, the MOAT. The predictive performance of the MOAT was fair (c-index 0.66), with a sensitivity of 90% for a 'medium-risk' category (specificity 30%), and 66% for a 'high-risk' category (specificity 61%). Decision curve analysis suggests that the MOAT has the potential to be clinically useful across a wide range of predicted risk probabilities (from approximately 15% to 70%).

Comparison with previous literature

Seven of the prediction tool studies included in the literature review (chapter 3) involved development of a prognostic model for adverse medication-related outcomes during hospitalisation. Of these, three predict adverse drug reactions (ADRs)[41 42 43], two predict adverse drug events (ADEs)[44 45], one predicts medication errors (MEs)[47], and one predicts MRPs[52].

Predictive performance was reported for six of these studies[41 42 44 45 47 52]. Of these, McElnay *et al*[44] concluded their model had insufficient sensitivity to be a satisfactory predictor of ADEs, but the remaining five studies reported adequate discriminatory capacity (c-index of 0.70-0.78 following bootstrap or external validation). Only two of these five studies reported calibration following development[42 47]. No studies reported calibration following external validation; it is therefore not possible to judge the potential accuracy of predictions in new datasets. Risk groups were created in two studies[41 42], permitting sensitivity and specificity to be reported:

- GerontoNet ADR risk score[41]:
    - developmental dataset – sensitivity 68% and specificity 65%;
    - validation dataset – although sensitivity and specificity not reported, it was possible to calculate from data provided, giving a sensitivity of 91%, and specificity of 27%;
- Brighten Adverse Drug Reactions Risk (BADRI) model[42]:

- developmental dataset – sensitivity 80% and specificity 55%;
- validation dataset – sensitivity 84% and specificity 43%.

The prognostic models developed by Nguyen *et al*[47] and Urbina *et al*[52], designed to target patients at risk of MEs and MRPs respectively, are potentially the most similar to the present study in terms of outcome event (given that MEs are a subset of MRPs). Despite this, key differences exist in terms of the modelling strategies used, which make direct comparison with the present study difficult:

- target population – Nguyen included medical and surgical patients, and Urbina included medical, surgical and maternity patients (the present study includes only medical patients);
- choice of candidate predictors – Nguyen did not investigate laboratory results, diagnostic groups, or comorbidities, all of which are included in the present study;
- method of outcome identification – Urbina used a computerised warning system, impacting on the type of MRPs identified.

In summary, although the predictive performance of the MOAT is broadly comparable with the existing prognostic models, it is not possible to make direct comparisons due to differences in the type of outcome predicted, and/or methods used for development. It is also not possible to compare potential clinical usefulness, as the existing studies did not report decision curve analysis.

Interpretation

While the discriminative ability of a prognostic model is important, Steyerberg advises that 'it is not possible to indicate a minimum value for the c-index to make a model clinically useful'[70]. This is because the c-index alone does not consider the consequence of false positive or false negatives predictions. For example, a model with a 'good' c-index (for example 0.8) will not be clinically useful if all predictions are above or below the optimal decision threshold[70 201]. The use of decision curve analysis was therefore helpful in permitting assessment of the MOAT's potential value in clinical practice. While the MOAT has a modest c-index (0.66), its predictions span a wide range of probabilities (from 9% to 86%), with net benefit across a significant range (15% to 70%). This suggests the MOAT has the potential to be useful in clinical practice in terms of guiding decision making at both decision thresholds (25% and 35%). Furthermore, the creation of three risk groups (low, medium and high-risk), permits pharmacists to take account of workload capacity when prioritising patients, as

does the reporting of both the predicted risk probability and risk group for individual patients.

Strengths and limitations

A strength of the approach taken for model development was adherence with PROGRESS[53 55], TRIPOD[94], and CHARMS[60] recommendations. This has the potential to enhance the quality of the modelling strategy, reduce bias, and facilitate full and detailed reporting to permit the quality and relevance of the study to be adequately assessed. Other strengths include:

- inclusion of two study sites, potentially increasing the generalisability of the MOAT;
- use of shrinkage techniques, increasing the potential accuracy of predictions in new patients;
- use of clinical decision curve analysis, informing potential clinical usefulness;
- development of an electronic scoring system, which aims to simplify use.

Limitations include the following:

- the presence of missing data, and subsequent use of multiple imputation. While this is a potential source of bias, the comparison of key characteristics of patients with and without missing data, and sensitivity analysis, suggest data were MAR. Given this missingness assumption, multiple imputation was less likely to introduce selection bias than complete-case analysis, in addition to being statistically more efficient. The use of multiple imputation also impacted on a number of the model diagnostic checks. For example, it was necessary to use non-imputed data for the review of 'outlying observations', resulting in 9 (0.6%) of 1,503 admissions being excluded from the review;
- investigations into possible 'interactions' between candidate predictors were not performed (i.e. an assessment of whether candidate predictors have a different association with the outcome depending on the value of a third variable). While interactions may have been present, none were hypothesised *a priori*, and it is recognised that a thorough assessment of possible interactions during modelling increases the risk of overfitting. As a result, Steyerberg recommends considering interactions only in studies with relatively large sample sizes[70];
- the decision threshold between low and medium-risk patients gives a sensitivity of 90%, meaning that 10% of patients who experienced an MSP MRP would be incorrectly categorised as low-risk, therefore would not receive pharmacist review.

The potential clinical consequence of these false negative results will be assessed in chapter 10;

- to achieve 90% sensitivity the corresponding specificity is approximately 30%. An improvement in specificity leads to a reduction in sensitivity (i.e. the decision threshold for high-risk patients has a specificity of 61% and sensitivity of 66%). This is a consequence of the ability to accurately predict MSP MRPs, but also the relatively high prevalence (40.6%);
- the potential complexity of the MOAT, requiring pharmacists to input 14 pieces of data. While electronic scoring has the potential to simplify use, usability will be investigated as part of the assessment of the MOAT (chapter 10);
- the requirement for pharmacists to correctly interpret definitions for candidate predictors, for example how to count comorbidities and number of 'regular' medicines. Guidance was therefore integrated into the electronic scoring system, and usability assessed as part of the MOAT assessment (chapter 10).

Implications

Given the above limitations, subsequent research may be able to improve on the present study by using prospective collection for predictor data; as discussed in chapter 6, this may overcome the need to imputing missing values.

Further research is required to assess the potential clinical consequence of false negative results, and usability of the MOAT in clinical practice. These will be assessed in chapter 10.

9.5 Conclusion

An electronic decision aid, the MOAT, was developed to permit identification of patients at highest risk of MSP MRPs. The predictive performance of the MOAT was fair (c-index 0.66), with a sensitivity of 90% for the 'medium-risk' category and 66% for the 'high-risk' category. Decision curve analysis suggests the MOAT has the potential to be clinically useful across a wide range of predicted risk probabilities.

The next chapter will describe assessment of the usability of the MOAT in clinical practice. This will involve a review of the content validity, ease of use, potential workload implications, and the potential clinical risk associated with false negative predictions.

Chapter 10: Assessment of the MOAT's clinical credibility

10.1 Introduction

The principal aim of this research was to develop a decision aid, the Medicines Optimisation Assessment Tool (MOAT^{TM}), to assist in the identification of adult patients at highest risk of moderate or severe preventable medication related problems (MSP MRPs) during hospital admission. The intention is for the MOAT to have the potential to be adopted widely into clinical practice, which will require clinical credibility, accuracy, generalisability and clinical effectiveness in improving decision making and the associated patient outcomes[206].

As discussed in chapter 9, the MOAT has reasonable predictive performance (concordance index 0.66), and the potential to be useful in guiding clinical decisions across a wide range of predicted risk probabilities. Assessment of the MOAT's generalisability and clinical effectiveness will require external validation, plus impact and implementation studies (discussed further in chapter 11); the aim of the work presented in this chapter was therefore to assess the MOAT's clinically credibility. The objectives were to:

- obtain consensus views of practising pharmacy professionals on the clinical credibility and usability of the MOAT;
- quantify the workload implications related to routine use;
- assess the potential clinical risk related to use of the MOAT;
- identify implications of the above findings for future development and implementation.

10.2 Methods

The clinical credibility of a prediction tool is dependent on a number of factors including content validity, ease of use, acceptability of the time taken to use the tool, and acceptability of the false negative rate[98 207]. This assessment therefore involved three stages: use of a consensus method to harness the insights of pharmacy professionals regarding the perceived clinical credibility and usability of the MOAT, an assessment of the workload implications, and an assessment of the clinical implications of false negative predictions. Each is described in turn below; the results are described in section 10.3.

Participation was voluntary. Pharmacists and clinical pharmacy technicians from Hospital A were given a participant information sheet (Appendix A10.1) and invited to volunteer for the consensus group and/or workload assessment. Written consent was obtained from all participants (Appendix A10.2).

10.2.1 Consensus views on the MOAT

The nominal group technique (NGT) was used to obtain consensus views of practising pharmacy staff on the clinical credibility of the MOAT[98 207]. The panellists were asked to consider the following questions:

- does the MOAT demonstrate content validity? That is to say, would most clinicians consider that the choice of predictors is appropriate for the purpose of the prediction tool, that no obvious predictors are missing, and that individual predictors are appropriately grouped?;
- is the visual presentation of the MOAT reasonable?;
- does the MOAT have the potential to be 'usable' in clinical practice (related to simplicity of interpretation and time taken to apply the MOAT)?

Consensus was also obtained on the choice of decision thresholds for the creation of risk categories (described in section 9.2.9.2). This involved consideration of the sensitivity versus specificity of alternative decision thresholds, and the balance between false negative predictions and the number of patients requiring pharmacist review.

Seven participants were included in the nominal group (as this has been reported as the maximum recommended number[101]). Pharmacists and clinical pharmacy technicians were invited to volunteer as the MOAT has the potential to be used by both professional groups; to ensure adequate representation at least three participants were required to be from one or other group.

A standard NGT method was used[101 208], with myself acting as facilitator. This comprised two meetings during which panellists discussed the issues, rated the questions, and then rerated the questions following further discussion. Questions were rated using a nine-point Likert scale; each question was worded as a statement, and participants asked to indicate the extent to which they agreed or disagreed with each statement, with a score of one indicating total disagreement, and nine indicating total agreement. A nine-point scale was used to permit the responses to be categorised, with scores of one to three representing disagreement, scores of four to six

representing an equivocal response, and scores of seven to nine representing agreement[208]. Consensus was interpreted as all participants' scores falling within one of these predefined three-point regions[208].

The format of the meetings was as follows.

Meeting one:

1. the MOAT was demonstrated to the group (by myself), including a description of the developmental process, and instructions for use;
2. participants were asked to reflect on one of the questions (specified above), and write down their views;
3. each participant, in turn, was invited to contribute one comment (either a statement or clarifying question), which was recorded on a flipchart;
4. when no more comments were forthcoming, there was a group discussion to clarify and evaluate the question;
5. participants were then given a rating sheet (Appendix A10.3) and asked to 'score' the question. Participants were also invited to give brief reasons for their score. This stage was confidential, and as facilitator, I emphasised that participants did not have to agree;
6. steps two to five were repeated until all questions had been scored.

Following this meeting, the responses were collated, with agreement for each question summarised using the median score and interquartile range (IQR) to establish central tendency and variability, treating responses as ordinal data. Where two adjacent scores were circled by a panellist, the mid-point between scores was used. The first-round scores were used to develop personalised score sheets for the second meeting. These included the same statements, together with the individual participant's scoring, median score, IQR, range, and all free text comments.

Meeting two:

1. the collated results from meeting one were presented (by myself), and the personalised second-round score sheets distributed to participants. As the MOAT workload assessment (section 10.2.2) took place in the interval between consensus meetings, these results, together with feedback received on use of the MOAT, were also presented and discussed;
2. participants were asked to reflect on one of the questions, taking into account their own score from the first meeting and the collated results (including summary scores and free text comments), and write down their views;

3. each participant, in turn, was invited to contribute one comment, which was recorded on a flipchart;
4. when no more comments were forthcoming, there was a group discussion to clarify and evaluate the question;
5. participants were then asked to 'rescore' the question. Participants were also invited to give brief reasons for this score. As before, this stage was confidential, and participants advised they did not have to agree;
6. steps two to five were repeated until all questions had been rescored.

Two panellists were unable to attend the second meeting due to unforeseen circumstances. I therefore met with each individually and shared the collated results from meeting one. Each panellist was then asked to reflect on each question in turn and given the opportunity to discuss their views. I then shared the discussion points raised by the group. The two panellists then rescored each statement in turn as described above.

10.2.2 Workload implications

The workload implications regarding use of the MOAT were assessed by:

- analysing the original dataset to estimate the proportion of patients who would be expected to require review by a pharmacist based on the decision thresholds selected during the nominal group meetings (i.e. the proportion of patients who would 'screen positive');
- calculating the median time required to apply the MOAT. This involved using the MOAT for a sample of patients (without acting on the findings).

For the second part of this assessment, four volunteers reviewed five patients each and recorded the time taken to use the MOAT (i.e. to obtain the required data and calculate the risk probability for each patient). Four assessors were used to account for potential inter-assessor variability. Following brief training on use of the MOAT (by myself), the assessors timed their use of the MOAT, and recorded the results using a data collection form designed for this purpose (Appendix A10.4). The median, range and IQR were then calculated. The median was calculated (rather than the mean) due to the non-parametric distribution of data.

10.2.3 Clinical implication of false negative predictions

The sensitivity of the MOAT (for the decision threshold selected to categorise patients as either low or medium-risk) was 90%, meaning that 10% of patients who experienced

an MSP MRP were categorised as low-risk, and therefore may not be selected for review by a pharmacist. To assess potential clinical risk associated with these false negative predictions, I reviewed the MSP MRPs experienced by these patients to identify the type, number and severity of these 'missed events'. The median severity score and IQR of MSP MRPs experienced by patients who 'screened positive' was compared with patients who 'screened negative', as was the median number and IQR of MSP MRPs experienced by both groups. Medians were compared (rather than means) due to the non-parametric distribution of data. The Mann-Whitney U test was used to test for differences.

10.3 Results

10.3.1 Consensus views on the MOAT

The consensus panel comprised four pharmacists and three clinical pharmacy technicians. This included two newly qualified pharmacists (Agenda for Change band 6), one mid-grade specialist pharmacist (band 7), and one senior specialist pharmacist (band 8a). Two of the clinical pharmacy technicians were band 6; one was band 7. The two meetings were held eight days apart; the first took two hours, and the second lasted one and a half hours.

The median score for each consensus statement, IQR, and range of scores are shown in Table 53. This shows that following meeting two, the median response and IQR were within the 'agreement category' for all five statements (i.e. within the range of seven to nine). Overall consensus, defined as all scores falling within one of the predefined three-point regions, was achieved for four of the five statements; one panellist gave the statement 'the time taken to use the MOAT is reasonable' a score of six, representing an equivocal response.

Table 53 – Consensus scores of practising pharmacy staff on the clinical credibility of the MOAT

Consensus statement	Scores – meeting 1*			Scores – meeting 2*		
	Median	IQR	Range	Median	IQR	Range
The choice of risk factors is appropriate	7	7 - 8	7 - 8	8	8 - 8	7 - 9
The presentation of the MOAT is reasonable	8	8 - 8	8 - 9	8	8 - 8	7 - 9
The MOAT is simple to interpret	8	8 - 9	7 - 9	8	8 - 9	8 - 9
The time taken to use the MOAT is reasonable	6	5 - 6.5†	5 - 8	8	7 - 8	6 - 8
The proposed 'decision thresholds' for the creation of risk groups are appropriate	7	6.5† - 7	6 - 8	7	7 - 8	7 - 8

* A score of one indicates total disagreement, and nine indicates total agreement. A nine-point scale was used to permit the responses to be categorised; a score of one to three represents disagreement, scores of four to six represent an equivocal response, and scores of seven to nine represent agreement. Consensus was interpreted as all scores falling within one of these predefined three-point regions

† Score taken to be 6.5 as both 6 and 7 circled by panellist

IQR = interquartile range, MOAT = Medicines Optimisation Assessment Tool

The written comments given by panellists are shown in Appendix A10.5. Regarding the choice of risk factors, opinion was divided as to whether additional risk factors should be included in the MOAT irrespective of their statistical significance, in particular the

use of anticoagulants and anti-diabetic medication. Some panellists felt these should be added, others felt that 'the line needed to be drawn somewhere', with only significant factors included. There were also suggestions that it may be helpful to list which risk factors were considered but excluded, as this may 'makes others trust the score more'. Additionally, it was suggested that all potential risk factors be included in the MOAT score if / when it becomes fully automated.

While the presentation and interpretability of the MOAT were generally rated as 'reasonable' (median score of eight for both statements), the following comments were received (as verbal comments recorded on the flipchart during the nominal group meetings and/or written comments on rating sheets) regarding potential improvements that could be made:

- the MOAT scoring screen should remain 'visible' while other programmes are opened (e.g. prescribing system and laboratory results) to facilitate data entry;
- add the ability to increase the size of the MOAT scoring screen (if required);
- increase size / boldness of writing on the scoring screen;
- remove Excel background from the behind the scoring screen;
- add an option for alternative units for serum creatinine (i.e. mg/dl) as this may be how it is reported at other hospitals;
- remove 'calculate buttons' for renal function and risk probability (as shown in Figure 21), as these could be calculated automatically once data submitted;
- add a help button to describe the probability and risk categories / explain how they are calculated.

The following issues were also raised as points for consideration during further development of the MOAT:

- may be difficult to distinguish 'race' by simply observing a patient;
- need to check the MOAT colour scheme is suitable for users with colour blindness / make sure the 'green risk score' is sufficiently visible on the green background;
- patients categorised as 'low-risk' should still be reviewed by a member of the pharmacy team at some point during hospital stay, for example by a clinical pharmacy technician, even if simply to enquire if the patient needs help with their medicines.

During the first meeting some panellists commented that it was difficult to rate the statement 'the time it takes to use the MOAT is reasonable' without first using the tool,

or having data on the time taken. As can be seen in Table 53, once these data were available the median score increased from six (equivocal) to eight (agree).

Regarding the choice of decision thresholds for the MOAT, the group appreciated the potential benefit of providing guidance to users on an appropriate course of action, and were able to agree on suitable cut-offs. Although panellists agreed that risk categories were helpful, they suggested that the 'percentage probability' may be the most useful output as this would permit prioritisation within risk categories. While not included as a written comment, the panellists also discussed the possibility of using the MOAT to allocate workload within the team, with initial allocation based on a patient's risk category (i.e. higher risk patients flagged for review by more experienced / qualified team members), with a referral system used to escalate care if required. It was suggested that this may improve the use of skill mix within the team, and provide greater clarity on the roles and responsibilities of team members.

10.3.2 Workload implications

As discussed in section 9.3.6.2, the MOAT's sensitivity was 90% (using the decision threshold to categorise patients as either low or medium-risk). This equates to approximately 78% of patients 'screening positive', therefore requiring pharmacist review (Table 51). Alternatively this can be interpreted as meaning that the MOAT permits identification of the 22% of patients least likely to experience an MSP MRP. Using the higher decision threshold (separating medium and high-risk patients), the sensitivity was 66.2%, which equates to 50% of patients requiring review (Table 52).

The time taken to apply the MOAT was assessed by four volunteers: three pharmacists and one clinical pharmacy technician (one band 6 pharmacist, two band 7 pharmacists, and one band 6 clinical pharmacy technician). The median time required to obtain the required information and calculate the risk score was 2 minutes 18 seconds per patient (range 1 minute 28 seconds to 5 minutes 4 seconds; IQR 1 minute 41 seconds to 3 minutes 12 seconds).

A potential improvement for the MOAT was suggested by one volunteer. This was to move the two risk factors that require laboratory results (serum creatinine and white cell count) next to each other to prevent the need to 'skip ahead when using the tool'.

10.3.3 Clinical implication of false negative predictions

Sixty-one patients experienced an MSP MRP despite being categorised as 'low-risk' (Table 51). The predicted risk probability for these 'false negative' patients is given in Appendix A10.6, together with brief details of the MSP MRPs experienced. This shows that the patients' predicted probability of experiencing an MSP MRP ranged from 12.2% to 25.2% and the number of MSP MRPs experienced per patient ranged from one to five (median 1, IQR 1 to 2). The median severity rating of the MSP MRPs experienced by these patients was 3.25 (IQR 3.0 to 3.5). There were no obvious trends in the type of MSP MRPs identified.

The 'true positive' patients (i.e. those patients correctly identified by the MOAT as being at risk of experiencing an MSP MRP) had predicted probabilities of experiencing an MSP MRP between 25.3% and 86.2%. The number of MSP MRPs per patient ranged from one to ten (median 1, IQR 1 to 2), and the median severity rating was 3.25 (IQR 3.0 to 3.75).

Despite the median severity rating of MSP MRPs being the same for the false negative and true positive patients, there was weak evidence for a statistically significant difference ($p = 0.046$). Similarly, despite the median number of MSP MRPs experienced per patient being the same in both groups, there was strong evidence for a difference ($p = 0.0021$).

10.4 Discussion

Key findings

The research presented in this chapter suggests that the MOAT was perceived as clinically credible and usable by practising pharmacy professionals. Additionally, the workload implications, based on the time taken to apply the MOAT compared to potential time saved by deprioritising low-risk patients, was considered to be reasonable. The results also suggest that patients who experienced an MSP MRP despite being categorised as low-risk (false negatives) may experience fewer MSP MRPs that are of lower severity, compared to patients categorised as medium or high-risk. A number of potential improvements were also identified, together with additional points to consider prior to implementation of the MOAT.

Interpretation

This research highlights the value of using expert opinion to guide future development of the MOAT, with a number of suggestions offered by participants to further improve the presentation and interpretability of the MOAT. While the MOAT will require external validation, plus impact and implementation studies prior to routine use, this assessment suggests that the MOAT has clinical credibility, which provides evidence to justify further development.

Strengths and limitations

Strengths of the research presented in this chapter include the:

- use of a standardised method (the NGT) to obtain consensus opinions on the MOAT;
- inclusion of pharmacists and clinical pharmacy technicians in the assessments (with varying levels of seniority) so increasing diversity;
- use of objective methods to calculate the workload implications, and clinical implications of false negative predictions.

Limitations include:

- the potential impact of selection and volunteer bias caused by the recruitment of all participants from the same study site, and the use of volunteers rather than random selection; as a result the participants may not be representative of all pharmacy staff at the study site, or those at other hospitals. This was necessary for pragmatic reasons, and the impact may be countered by the diversity within the assessors in terms of profession role and seniority;

- a risk of moderator bias, with myself acting as facilitator for the NGT meetings; to address this I endeavoured to maintain an objective approach throughout, providing facts and clarification (where needed) rather than personal opinions. The questions posed during the NGT meetings were also selected from previous literature, with the aim of minimising the potential that question bias may influence panellists' responses. Reporting bias was minimised by verbatim reporting of panellists' written comments.

Implications for development of the MOAT and future research

It will be necessary to repeat the clinical credibility and workload assessments following external validation of the MOAT to take account of predictive accuracy in a new group of patients. Given the limitations identified above, ideally this should involve a more diverse group of pharmacy staff, and an impartial facilitator.

The work presented in this chapter has provided valuable user-feedback, guiding future development of the MOAT in terms of its presentation and usability (as detailed in sections 10.3.1 and 10.3.2). Opinions varied on the inclusion of additional risk factors, specifically anticoagulants and anti-diabetic medication; work will therefore be required to investigate this further prior to implementation (on the basis that inclusion would have minimal impact on patients' risk scores, but may increase clinical confidence in the MOAT).

This initial workload assessment suggests the MOAT has the potential to increase the efficiency of clinical pharmacy services, but this will depend on how the MOAT is applied in practice. For example, if scored manually (using the electronic scoring system) it may not be feasible, or beneficial, to rescore every patient every day. Research will be required to investigate this further, but one possibility would be to 'MOAT score' patients on hospital admission to determine the level of review required, for example to indicate if medicines reconciliation is required, and/or to allocate team members appropriately based on their knowledge and skills. Subsequent prioritisation decisions could then be guided by professional judgement, or MOAT scores could be recalculated if there is a significant change in risk, for example due to the initiation of high-risk medicines, resulting in escalation or de-escalation as appropriate.

Further research may also permit potential efficiency savings to be quantified. For example, if the MOAT were used to select patients for pharmacy-led medicines reconciliation (with medicines reconciliation undertaken for medium and high-risk patients only) it may be possible to save approximately 12.5 minutes of pharmacy staff

time for all low-risk patients (based on the estimate that a medicines reconciliation takes an average of 15 minutes[209], and the median time to apply the MOAT is 2 minutes 18 seconds). If applied to all low-risk patients (i.e. 22% of admissions), this may lead to a significant and positive impact on workload.

Further work will also be required to establish the clinical acceptability of this type of approach to prioritisation. The nominal group members expressed some unease over the potential for patients to be 'overlooked' by pharmacy services, as illustrated by the comment 'don't wish to be in acute [the acute assessment wards] when all low-risk patients are not seen at all by pharmacy'. One possibility, as discussed by the group, may be to develop guidelines on the level of pharmacy input required dependent on risk categorisation. This could range from a simple face-to-face discussion with low-risk patients to more intensive interventions for patients in higher risk categories (such as medicines reconciliation and medication review). As discussed in chapter 5, it may also be possible to combine the MOAT with other triggers for pharmacy review; for example, patients with swallowing difficulties, those receiving end of life care, or those at risk of MRPs post discharge. Potentially the MOAT could then be used by ward-based pharmacy staff as part of a suite of tools, permitting prioritisation of patients, and appropriate allocation of workload between team members based on their skills and expertise. The development of these types of implementation strategies may also address patients' views of safety; the medical view of patient safety often focusses on outcomes and avoidance of harm[210], whereas patients tend to focus on what makes them 'feel safe', including processes of care, and interpersonal dynamics with care providers[211 212]. The MOAT inherently fits a medical view of safety, with attention on 'risk reduction'; incorporation of the MOAT into an holistic system offering some level of pharmacy input to all patients, including those categorised as 'low-risk', may therefore help provide a sense of safety for all patients.

Clinical acceptability may also be dependent on the choice of decision thresholds used to categorise patients, as this will impact on both workload and the false negative rate; acceptance may therefore depend on workforce capacity and aversion to risk within organisations or clinical teams. Research may therefore be warranted to explore this further, with the potential to develop flexible thresholds that could be tailored either to organisational need, or to fluctuating staffing levels.

Finally, an additional area of potential future research would be the development of an automated scoring system that is fully integrated into the relevant electronic data

sources. This would permit risk assessments in 'real-time', further supporting implementation.

10.5 Conclusion

Standardised and objective methods were used to assess the MOAT's clinical credibility, and to identify implications for future development and implementation. The results suggest that the MOAT is clinically credible, minor modifications were also identified that have the potential to further improve usability. Further research will be required to establish generalisability and clinical effectiveness.

The next chapter will summarise the research presented in this thesis by way of an overall discussion.

Chapter 11: Overall discussion

Since starting the research presented in this thesis, medication safety has continued to be high on international and national agendas. In 2017 the World Health Organization published their Global Patient Safety Challenge: Medication Without Harm[7], which outlines their global initiative to reduce the level of severe, avoidable medication-related harm by 50% over five years. This was driven by the recognition that 'unsafe medication practices and medication errors are a leading cause of injury and avoidable harm in health care systems across the world', with an estimated global cost of $42 billion United States dollars annually[7]. In England, the Department of Health and Social Care commissioned a review of the prevalence and economic burden of medication errors in the English NHS, resulting in publication of a report by the Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) in February 2018[213]. This report received ministerial attention, informing an article in the Daily Telegraph by the Secretary of State Jeremy Hunt, and resulting in significant media interest[214]; national headlines included 'NHS medication errors contribute to as many as 22,000 deaths a year'[215], and 'Drug mistakes killing up to 22,300 patients a year'[216]. While the majority of medicine use occurs in primary care, the safe use of medicines in secondary care, and at transitions of care, continue to be areas of attention[7 213], together with ongoing calls for hospital pharmacy services to operate more efficiently and safely[36]. The work presented in this thesis therefore remains as topical and relevant as when first proposed in 2015.

A summary of the key findings of this research are described below, together with a high level comparison with previous literature, a summary of the overall strengths and limitations, and the implications for practice and for further research.

11.1 Summary of key findings

The aim of this research was to develop a prediction tool, the Medicines Optimisation Assessment Tool (MOAT^{TM}), to target patients most in need of pharmacists' input while in hospital; this was driven by a desire to increase the efficiency of hospital pharmacy services, reduce risks and improve patient outcomes.

One hundred and eighteen potential prognostic factors, also known as candidate predictors, were identified from previous research and an expert survey, and 18 pre-selected for MOAT development. Among 1,503 eligible patient admissions, 894 (59.5%) experienced at least one medication related problem (MRP), with 610 (40.6%)

experiencing the study's outcome event, namely at least one moderate or severe preventable MRP (MSP MRP). A prognostic model was developed using multivariable logistic regression to determine the relationship between candidate predictors and patients with an outcome event, and used to develop an electronic clinical decision aid, the MOAT. The predictive performance of the MOAT was fair (concordance index 0.66), and calibration was good. Three risk groups were created to categorise patients as low, medium or high-risk of experiencing an MSP MRP. The decision threshold for 'low' and 'medium-risk' patients has a sensitivity of 90% (specificity 30%); the sensitivity for the threshold between 'medium' and 'high-risk' patients is 66% (specificity 61%). Decision curve analysis suggests that the MOAT has potential to be clinically useful across a wide range of predicted risk probabilities (from approximately 15% to 70%). The MOAT was assessed in terms of content validity, ease of use, acceptability of the time taken to use the tool, and acceptability of the false negative rate; results suggest the MOAT is clinically credible, and has potential to increase the efficiency of hospital pharmacy services by identifying the 22% of patients least likely to experience an MSP MRP.

11.2 Comparison with previous literature

The literature review performed for this study (chapter 3) found that no methodologically sound prognostic model to target hospital patients based on their risk of MSP MRPs currently exists. Twelve prediction tools for adverse medication-related outcomes were reviewed; five were developed using consensus methods[46 48-51], and of the seven statistically derived tools, three predict adverse drug reactions (ADRs)[41 42 43], two predict adverse drug events (ADEs)[44 45], one predicts medication errors (MEs)[47], and one predicts MRPs[52]. The prognostic models developed by Nguyen *et al*[47] and Urbina *et al*[52], designed to target patients at risk of MEs and MRPs respectively, are potentially the most similar to the present study in terms of outcome event (given that MEs are a subset of MRPs). Despite this, key differences exist in terms of the modelling strategies used, which limit direct comparison with the present study. Clinical usefulness, assessed using decision curve analysis, was not reported for any of the existing statistically derived tools.

The MRP prevalence for the present study was found to be 59.5%, which is consistent with previous research (52% to 81%[72 78 90]). While no estimation of the prevalence of MSP MRPs exists, Blix *et al*[78] reported that 49.6% of MRPs identified during their study were 'extremely important or major'. Although the severity grading system used by Blix *et al* is not directly comparable with the present study, the proportion of MRPs that may

be considered to be 'clinically significant' is broadly comparable, with 44.2% of MOAT MRPs being found to be 'moderate or severe' (overall prevalence of MSP MRPs 40.6%).

11.3 Strengths and limitations

A strength of this research has been adherence with recommendations of the PROGnosis RESearch Strategy (PROGRESS) partnership[53 55], Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement[94], and CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS)[60], at all stages of MOAT development, from initial study design to statistical analysis and assessment of model performance. Other strengths include the:

- inclusion of two study sites with markedly different patient demographics, potentially increasing the generalisability of the MOAT;
- use of MSP MRPs as the outcome event for the MOAT, so aligning the MOAT with pharmaceutical care practice, enhancing clinical credibility, clinical relevance, and feasibility of implementation in terms of workload for pharmacists;
- inclusion of 1,503 study admissions with 610 outcome events, resulting in 18 'events per variable' (EPV), which is significantly above the 'rule of thumb' of at least 10 EPV (to address the risk of model overfitting[92]);
- robustness of definitions and data collection procedures for the outcome event and candidate predictors, increasing reliability and reproducibility[94];
- use of internal validation and adjustment for optimism, improving potential predictive accuracy of the MOAT when used in a new group of patients[70 94];
- development of an electronic decision aid to simplify use, and indicate a course of action[70 94];
- creation of two decision thresholds for the MOAT, permitting flexible prioritisation based on clinical risk and workforce capacity;
- use of 'expert opinion' to inform selection of candidate predictors, validate and severity rate MRPs, select decision thresholds for the MOAT's risk categories, and assess clinical credibility of the MOAT.

A limitation of the study is the possible underestimation of the prevalence of MRPs. As discussed in chapter 7, pharmacists at the study sites identified approximately 85% of MRPs as part of a simulated MRP identification assessment exercise. While a percentage agreement of 80% may be considered acceptable in health research[164],

identification of less than 100% of outcome events may be subject to criticism, highlighting the need for robust external validation of the MOAT, including the possible need for updating or recalibratation[94].

Other limitations include the:

- exclusion of predictors that are not routinely measured / recorded in clinical practice, have low prevalence, or potential measurement error (due to the potential for inaccurate results[60 63 95]). For example, it was not possible to model 'non-compliance with medication' as this is not routinely assessed or recorded, and has potential for measurement error. Similarly, high-risk medicines, such as cytotoxics, could not be included as an individual category due to infrequent use. While it may be appropriate to omit predictors if their effect cannot be reliably estimated[70], data on which predictors were excluded from the analysis will need to be shared with MOAT users to inform implementation;
- observational nature of the study, meaning data collection was not carried out under strict trial conditions. While this may have impacted on the robustness of data collection, it permitted the MOAT to reflect clinical practice in terms of MRP identification, and ensured use of routinely recorded predictor data;
- presence of missing predictor data, and subsequent use of multiple imputation. While this is a potential source of bias, a sensitivity analysis suggested data were missing at random (MAR), therefore multiple imputation was less likely to introduce selection bias than complete-case analysis[94 148], in addition to being statistically more efficient[62];
- use of simplified categorisation of predictors. This prevented detailed analysis (for example of individual classes of medicines such as insulins) but was necessary to reduce the risk of overfitting (associated with the use of high numbers of variables in model development[60 62 94]);
- trade-off between sensitivity and specificity of the MOAT[217]. To achieve a sensitivity of 90% (i.e. the decision threshold between 'low' and 'medium-risk' patients) the MOAT's has a relatively low specificity (30%). The specificity for the threshold between 'medium' and 'high-risk' patients is higher (61%), but at the expense of lower sensitivity (66%). While no predictive model is perfectly accurate[217], further work will be required to investigate the acceptability of the MOAT's false negative rate, and the impact on clinical confidence;
- need for external validation to establish generalisability, both within the United Kingdom and more widely.

11.4 Interpretation

This research was successful in meeting it proposed objectives, culminating in development of the MOAT, and assessment of its predictive performance and clinical credibility. To the best of my knowledge, the MOAT is the first evidence-based clinical prioritisation tool to identify patients most in need of pharmacists' input in terms of their risk of MSP MRPs. Results suggest the predictive accuracy, clinical usefulness and clinical credibility of the MOAT are acceptable, providing evidence to justify further development.

Given the above limitations, subsequent research may be able to improve on the present study through use of alternative data collection methods that increase the identification rate for MRPs, and reduce the occurrence of missing predictor data. The inclusion of a larger study sample may also permit use of more complex categorisation of predictors.

11.5 Implications for practice

It is not possible to advocate routine use of the MOAT prior to completion of external validation; impact and implementation studies will also be needed[206]. Subject to completion of this work, the MOAT has potential to be applicable to adult hospitalised medical patients (general, acute, and elderly medicine), irrespective of age. Given the diverse characteristics of the sample population (section 6.4.2), the MOAT has potential applicability to a wide range of patients in terms of age, ethnicity, comorbidities and presenting medical conditions.

As discussed in chapter 10, introduction of the MOAT may require the development of implementation strategies regarding the level of pharmacy input required by patients dependent on their risk categorisation. This may have implications for hospital pharmacy managers and others, who will need to consider the approach to risk management and governance that may be needed.

11.6 Implications for further research

Extensive external validation, involving prospective validation in a new cohort, will be required to further assess accuracy and generalisability before routine use of the MOAT could be recommended[55]. External validation will also provide opportunity to refine the MOAT in terms of improving the accuracy such as by updating the model[55] [218], and/or simplifying the scoring system.

Following external validation, impact and implementation studies will be required to establish whether the MOAT has advantages over current practice, is compatible with (and can easily be incorporated into) practice, has the potential to change pharmacists' behaviour, has a positive impact on patient outcomes, and is cost effective[206]. In terms of compatibility with practice, various potential implementation strategies were discussed in chapter 10. These include use of the MOAT to determine the level of pharmacy review required by individual patients, and combining the MOAT with other triggers for pharmacy review, for example, swallowing difficulties, end of life care, or risk of MRPs post discharge; potentially the MOAT could then be used as part of a suite of tools, permitting prioritisation of patients, and appropriate allocation of workload between team members based on skills and expertise.

Other potential future developments for the MOAT include:

- development of separate risk prediction tools that are specific to different stages of a 'patient stay' in hospital (as discussed in chapter 8);
- integration of the MOAT into automated systems such as electronic health records systems. This could result in the ability to perform accurate, automated risk assessments in 'real-time', which would further support implementation into clinical environments;
- assessment of the transportability of the MOAT (i.e. the ability to produce accurate predictions among people drawn from different but plausibly related populations[219], such as surgical patients, or patients in care homes). This may require model adjustment, or development of new prognostic models.

11.7 Overall conclusion

Overall this thesis has extended current knowledge in the field of clinical prioritisation through development of a methodologically sound prognostic model, the MOAT, to target hospital patients most in need of pharmacists' input. To my knowledge, this is the first prognostic model to identify hospitalised medical patients at risk of MSP MRPs.

In addition, this research has created knowledge that may inform future studies in this field, including:

- identification of potential risk factors associated with the occurrence of MRPs in hospitalised patients, using both theoretical knowledge and expert opinion;
- the prevalence of MRPs and MSP MRPs in hospitalised UK patients;

- quantification of the potential variability in MRP identification by hospital pharmacists;
- consensus views of practising pharmacy clinicians on the requirements of a predictive tool, including presentation and usability.

Extensive external validation, involving prospective validation in a new cohort, will be required to further assess accuracy and generalisability of the MOAT before routine use can be recommended. Further research will also be required in terms of impact and implementation studies to assess the extent to which use of the MOAT may affect decision making, improve efficiency or improve health outcomes.

References

1. Francis R. Report of the Mid Staffordshire NHS Foundation Trust Public Enquiry: Mid Staffordshire NHS Foundation Trust Public Inquiry 2013, 2013.
2. Berwick D. A promise to learn – a commitment to act. Improving the safety of patients in England. In: England TNAGotSoPi, ed., 2013.
3. The Royal Pharmaceutical Society. Keeping patients safe when they transfer between care providers – getting the medicines right, 2012.
4. The Royal Pharmaceutical Society. Medicines Optimisation: Helping patients to make the most of medicines, good practice guidance for healthcare professionals in England, 2013.
5. National Institute for Health and Care Excellence. Medicines optimisation: the safe and effective use of medicines to enable the best possible outcomes, NICE guidelines [NG5], 2015.
6. National Institute for Health and Care Excellence. CG138 Patient experience in adult NHS services, 2012.
7. World Health Organization. WHO Global Patient Safety Challenge: Medication Without Harm, 2017.
8. Otero M-j, Schmitt E. Clarifying terminology for adverse drug events. *Annals of internal medicine* 2005;142(1):77; author reply 77.
9. Avery AA, Barber N, Ghaleb M, et al. Investigating the prevalence and causes of prescribing errors in general practice: the PRACtICe study. 2012
10. Pirmohamed M, James S, Meakin S, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18 820 patients. *BMJ* 2004;329(7456):15-19. doi: 10.1136/bmj.329.7456.15
11. Dornan T, Ashcroft D, Heathfield H, et al. An in-depth investigation into causes of prescribing errors by foundation trainees in relation to their medical education: EQUIP study. *London: General Medical Council* 2009:1-215.
12. van den Bemt PMLA, Egberts TCG, de Jong-van den Berg LTW, et al. Drug-Related Problems in Hospitalised Patients. *Drug Safety* 2000;22(4):321-33. doi: 10.2165/00002018-200022040-00005
13. NHS England/MHRA. Patient Safety Alert. Stage Three: Directive. Improving medication error incident reporting and learning, 2014.
14. Frontier Economics Ltd. Exploring the costs of unsafe care in the NHS. A REPORT PREPARED FOR THE DEPARTMENT OF HEALTH. 2014

15. Karnon J, McIntosh A, Dean J, et al. Modelling the expected net benefits of interventions to reduce the burden of medication errors. *Journal of health services research & policy* 2008;13(2):85-91.
16. Krähenbühl-Melcher A, Schlienger R, Lampert M, et al. Drug-Related Problems in Hospitals. *Drug Safety* 2007;30(5):379-407. doi: 10.2165/00002018-200730050-00003
17. Garfield S, Barber N, Walley P, et al. Quality of medication use in primary care - mapping the problem, working to a solution: a systematic review of the literature. *BMC Medicine* 2009;7 doi: 10.1186/1741-7015-7-50 [published Online First: 21 September 2009]
18. Lund BC. Adverse drug events in older adults: will risk factor algorithms translate into effective clinical interventions? *Expert Review of Clinical Pharmacology* 2011;4(6):655-57. doi: http://dx.doi.org.libproxy.ucl.ac.uk/10.1586/ecp.11.48
19. Committee of Experts on Management of Safety and Quality in Health Care (SP-SQS) Expert Group on Safe Medication Practices. Glossary of terms related to patient and medication safety 2005 [Available from: http://www.bvs.org.ar/pdf/seguridadpaciente.pdf.
20. Pharmaceutical Care Network Europe. The PCNE Classification V 7.0 2016 [Available from: http://www.pcne.org/upload/files/152_PCNE_classification_V7-0.pdf accessed March 2017.
21. The Royal Pharmaceutical Society. Professional Standards for Hospital Pharmacy Services, version 3, 2017.
22. Stephens M. Hospital Pharmacy. Second edition ed: Pharmaceutical Press 2011.
23. The Society of Hospital Pharmacists of Australia. Standards of Practice for Clinical Pharmacy Services, 2016.
24. American College of Clinical Pharmacy. Definition of Clinical Pharmacy [Available from: https://www.accp.com/stunet/compass/definition.aspx accessed March 2017.
25. Kaboli PJ, Hoth AB, McClimon BJ, et al. Clinical pharmacists and inpatient medical care: a systematic review. *Archives of internal medicine* 2006;166(9):955-64.
26. Chisholm-Burns MA, Lee JK, Spivey CA, et al. US pharmacists' effect as team members on patient care: systematic review and meta-analyses. *Medical care* 2010;48(10):923-33.
27. The Royal Pharmaceutical Society. Now or Never: Shaping pharmacy for the future, 2013.
28. NHS. Five Year Forward View, 2014.

29. Lord Carter of Coles. Operational productivity and performance in English NHS acute hospitals: Unwarranted variations, 2016.
30. Nuffield Trust. A decade of austerity? The funding pressures facing the NHS from 2010/11 to 2021/22, 2012.
31. The King's Fund. Understanding NHS financial pressures - How are they affecting patient care? , 2017.
32. East & South East England Specialist Pharmacy Services. Prioritising pharmaceutical care delivery at ward level – Vs.1 2011.
33. Health Quality and Safety Commission New Zealand. All hands on deck: prioritisation criteria 2011 [Available from: https://www.hqsc.govt.nz/assets/Medication-Safety/Med-Rec-PR/MR-Workshop-2011/MR-Workshop-All-hands-on-deck-Prioritisation-criteria-Nirasha-Parsotam.pdf accessed March 2017.
34. American Society of Health-System Pharmacists. The consensus of the Pharmacy Practice Model Summit. *American Journal of Health-System Pharmacy* 2011;68:1148-52.
35. Dodds LJ. Optimising pharmacy input to medicines reconciliation at admission to hospital: lessons from a collaborative service evaluation of pharmacy-led medicines reconciliation services in 30 acute hospitals in England. *European Journal of Hospital Pharmacy* 2014;21(2):95-101. doi: 10.1136/ejhpharm-2013-000385
36. NHS England. Transformation of seven day clinical pharmacy services in acute hospitals, 2016.
37. Moore A. Standardise, upskill and scale up: how one acute trust is facing the Carter challenge. *The Pharmaceutical Journal* 2016;297(7894):205-07.
38. Kaufmann CP, Stämpfli D, Hersberger KE, et al. Determination of risk factors for drug-related problems: a multidisciplinary triangulation process. *BMJ Open* 2015;5(3) doi: 10.1136/bmjopen-2014-006376
39. Joint British Societies recommendations on the prevention of Cardiovascular Disease. Risk Calculator [Available from: http://www.jbs3risk.com/pages/risk_calculator.htm accessed March 2017.
40. Waterlow J. Waterlow Pressure Ulcer Prevention/Treatment Policy 2005 [Available from: http://www.judy-waterlow.co.uk/downloads/Waterlow%20Score%20Card-front.pdf accessed March 2017.
41. Onder G, Petrovic M, Tangiisuran B, et al. Development and validation of a score to assess risk of adverse drug reactions among in-hospital patients 65 years or

older: the GerontoNet ADR risk score. *Archives of internal medicine* 2010;170(13):1142-48.

42. Tangiisuran B, Scutt G, Stevenson J, et al. Development and validation of a risk model for predicting adverse drug reactions in older people during hospital stay: Brighton Adverse Drug Reactions Risk (BADRI) model. *PloS one* 2014;9(10):e111254.
43. Kiguba R, Karamagi C, Bird SM. Incidence, risk factors and risk prediction of hospital-acquired suspected adverse drug reactions: a prospective cohort of Ugandan inpatients. *BMJ Open* 2017;7(1) doi: 10.1136/bmjopen-2015-010568
44. McElnay J, McCallion C, Al-Deagi F, et al. Development of a risk model for adverse drug events in the elderly. *Clinical drug investigation* 1997;13(1):47-55.
45. Trivalle C, Burlaud A, Ducimetière P, et al. Risk factors for adverse drug events in hospitalized elderly patients: a geriatric score. *European Geriatric Medicine* 2011;2(5):284-89.
46. Saedder EA, Lisby M, Nielsen LP, et al. Detection of patients at high risk of medication errors: development and validation of an algorithm. *Basic & clinical pharmacology & toxicology* 2016;118(2):143-49.
47. Nguyen T-L, Leguelinel-Blache G, Kinowski J-M, et al. Improving medication safety: Development and impact of a multivariate model-based strategy to target high-risk patients. *PloS one* 2017;12(2):e0171995.
48. Cottrell R, Caldwell M, Jardine G. Developing and implementing a pharmacy risk screening tool. *Hospital Pharmacy Europe* 2013(71):58-60.
49. Falconer N, Nand S, Liow D, et al. Development of an electronic patient prioritization tool for clinical pharmacist interventions. *American Journal of Health-System Pharmacy* 2014;71(4):311-20. doi: 10.2146/ajhp130247
50. Roten I, Marty S, Beney J. Electronic screening of medical records to detect inpatients at risk of drug-related problems. *Pharmacy World & Science* 2009;32(1):103. doi: 10.1007/s11096-009-9352-6
51. Hickson RP, Steinke DT, Skitterall C, et al. Evaluation of a pharmaceutical assessment screening tool to measure patient acuity and prioritise pharmaceutical care in a UK hospital. *European Journal of Hospital Pharmacy* 2016 doi: 10.1136/ejhpharm-2015-000829
52. Urbina O, Ferrández O, Grau S, et al. Design of a score to identify hospitalized patients at risk of drug-related problems. *Pharmacoepidemiology and drug safety* 2014;23(9):923-32.

53. Hemingway H, Croft P, Perel P, et al. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *Bmj* 2013;346:e5595.
54. Peat G, Riley RD, Croft P, et al. Improving the transparency of prognosis research: the role of reporting, data sharing, registration, and protocols. *PLoS Med* 2014;11(7):e1001671.
55. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10(2):e1001381.
56. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012;9(5):e1001221.
57. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): The tripod statement. *Annals of Internal Medicine* 2015;162(1):55-63. doi: 10.7326/M14-0697
58. Riley RD, Hayden JA, Steyerberg EW, et al. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;10(2):e1001380.
59. Hingorani AD, van der Windt DA, Riley RD, et al. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *Bmj* 2013;346:e5793.
60. Moons KG, de Groot JA, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;11(10):e1001744.
61. Stevenson JM, Williams JL, Burnham TG, et al. Predicting adverse drug reactions in older adults; a systematic review of the risk prediction models. *Clinical interventions in aging* 2014;9:1581.
62. Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: developing a prognostic model. *Bmj* 2009;338:b604.
63. Katz MH. Multivariable analysis: a primer for readers of medical research. *Annals of internal medicine* 2003;138(8):644-50.
64. Suggett E, Marriott J. Risk Factors Associated with the Requirement for Pharmaceutical Intervention in the Hospital Setting: A Systematic Review of the Literature. *Drugs - Real World Outcomes* 2016;3(3):241-63. doi: 10.1007/s40801-016-0083-4
65. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC medical research methodology* 2007;7(1):10.

66. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6(7):e1000100.
67. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS med* 2009;6(7):e1000097.
68. MacLure K, Paudyal V, Stewart D. Reviewing the literature, how systematic is systematic? *International journal of clinical pharmacy* 2016;38(3):685-94.
69. Hayden JA, van der Windt DA, Cartwright JL, et al. ASsessing bias in studies of prognostic factors. *Annals of Internal Medicine* 2013;158(4):280-86. doi: 10.7326/0003-4819-158-4-201302190-00009
70. Steyerberg E. Clinical prediction models : a practical approach to development, validation and updating.
71. Morrison C. Improving patient safety through changing a clinical pharmacy service. *The Pharmaceutical Journal* 2014;292(7806/7):426. doi: 10.1211/PJ.2014.11137445
72. Ayalew MB, Megersa TN, Mengistu YT. Drug-related problems in medical wards of Tikur Anbessa specialized hospital, Ethiopia. *Journal of research in pharmacy practice* 2015;4(4):216.
73. Lenssen R, Heidenreich A, Schulz JB, et al. Analysis of drug-related problems in three departments of a German University hospital. *International journal of clinical pharmacy* 2016;38(1):119-26.
74. Allen JE. Risk factors and prevention strategies for adverse drug reactions. *Journal of Pharmaceutical Care in Pain & Symptom Control* 1995;3(1):31-35.
75. MacKinnon NJ, Helper CD. Indicators of preventable drug-related morbidity in older adults. *Journal of Managed Care Pharmacy* 2003;9(2):134-41.
76. Bates DW, Miller EB, Cullen DJ, et al. Patient risk factors for adverse drug events in hospitalized patients. *Archives of internal medicine* 1999;159(21):2553-60.
77. Van den Bemt P, Egberts A, Lenderink A, et al. Risk factors for the development of adverse drug events in hospitalized patients. *Pharmacy World & Science* 2000;22(2):62-66.
78. Blix HS, Viktil KK, Reikvam Å, et al. The majority of hospitalised patients have drug-related problems: results from a prospective study in general hospitals. *European journal of clinical pharmacology* 2004;60(9):651-58.
79. Evans RS, Lloyd JF, Stoddard GJ, et al. Risk factors for adverse drug events: a 10-year analysis. *Annals of Pharmacotherapy* 2005;39(7-8):1161-68.

80. Johnston PE, France DJ, Byrne DW, et al. Assessment of adverse drug events among patients in a tertiary care medical center. *American journal of health-system pharmacy* 2006;63(22):2218-27.
81. Zopf Y, Rabe C, Neubert A, et al. Risk factors associated with adverse drug reactions following hospital admission: a prospective analysis of 907 patients in two German university hospitals. *Drug Safety* 2008;31(9):789-99.
82. Davies EC, Green CF, Taylor S, et al. Adverse drug reactions in hospital in-patients: a prospective analysis of 3695 patient-episodes. *PLoS one* 2009;4(2):e4439.
83. Zaal RJ, van Doormaal JE, Lenderink AW, et al. Comparison of potential risk factors for medication errors with and without patient harm. *Pharmacoepidemiology and drug safety* 2010;19(8):825-33.
84. Muñoz-Torrero JFS, Barquilla P, Velasco R, et al. Adverse drug reactions in internal medicine units and associated risk factors. *European journal of clinical pharmacology* 2010;66(12):1257-64.
85. Dequito AB, Mol PG, van Doormaal JE, et al. Preventable and non-preventable adverse drug events in hospitalized patients. *Drug safety* 2011;34(11):1089.
86. Ben-Yehuda A, Bitton Y, Sharon P, et al. Risk factors for prescribing and transcribing medication errors among elderly patients during acute hospitalization. *Drugs & aging* 2011;28(6):491.
87. Beckett RD, Sheehan AH, Reddan JG. Factors associated with reported preventable adverse drug events: a retrospective, case-control study. *Annals of Pharmacotherapy* 2012;46(5):634-41.
88. O'connor MN, Gallagher P, Byrne S, et al. Adverse drug reactions in older patients during hospitalisation: are they predictable? *Age and ageing* 2012;41(6):771-76.
89. Sikdar KC, Dowden J, Alaghehbandan R, et al. Adverse drug reactions in elderly hospitalized patients: a 12-year population-based retrospective cohort study. *Annals of Pharmacotherapy* 2012;46(7-8):960-71.
90. Wilmer CM, Huiskes VJB, Natsch S, et al. Drug-related problems in a clinical setting: a literature review and cross-sectional study evaluating factors to identify patients at risk. *Eur J Hosp Pharm* 2015;22(4):229-35.
91. Ashcroft DM, Lewis PJ, Tully MP, et al. Prevalence, nature, severity and risk factors for prescribing errors in hospital inpatients: prospective study in 20 UK hospitals. *Drug safety* 2015;38(9):833-43.

92. Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per variable in logistic regression analysis. *Journal of clinical epidemiology* 1996;49(12):1373-79.
93. Moons KG, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? *Bmj* 2009;338:b375.
94. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and ElaborationThe TRIPOD Statement: Explanation and Elaboration. *Annals of internal medicine* 2015;162(1):W1-W73.
95. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *American journal of epidemiology* 2007;165(6):710-18.
96. Harrell F. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis: Springer 2015.
97. Medina LS, Zurakowski D. Measurement variability and confidence intervals in medicine: why should radiologists care? *Radiology* 2003;226(2):297-301.
98. Laupacis A, Sekar N. Clinical prediction rules: a review and suggested modifications of methodological standards. *Jama* 1997;277(6):488-94.
99. Moons KG, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients?: Pergamon, 2002.
100. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Statistics in medicine* 2004;23(10):1631-60.
101. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. *International journal of clinical pharmacy* 2016;38(3):655-62.
102. Lee Y-h, Bang H, Kim DJ. How to establish clinical prediction models. *Endocrinology and Metabolism* 2016;31(1):38-44.
103. Petrovic M, Tangiisuran B, Rajkumar C, et al. Predicting the risk of adverse drug reactions in older inpatients: external validation of the GerontoNet ADR Risk Score using the CRIME cohort. *Drugs & aging* 2017;34(2):135-42.
104. Stevenson J, Kindsiko K, Ikpemo C, et al. Assessing ADR risk in older patients: Do the prediction models agree? *Age and Ageing* 2013;42(suppl_3):iii26.
105. Falconer N, Liow D, Zeng I, et al. Validation of the assessment of risk tool: patient prioritisation technology for clinical pharmacist interventions. *Eur J Hosp Pharm* 2017:ejhpharm-2016-001165.

106. Bonnerup DK, Lisby M, Sædder EA, et al. Risk of prescribing errors in acutely admitted patients: a pilot study. *International journal of clinical pharmacy* 2016;38(5):1157-63.
107. Lewis P. Developing and assessing the feasibility of a pharmaceutical care complexity screening tool to aid the targeted delivery of patient focussed clinical pharmacy services in hospital: Europe PMC plus; [Available from: https://europepmc.org/grantfinder/grantdetails?query=pi:%22Lewis+P%22+gid:%22PB-PG-1215-20031%22+ga:%22DH/NIHR%22 accessed July 2017.
108. George J, Phun Y-T, Bailey MJ, et al. Development and validation of the medication regimen complexity index. *Annals of Pharmacotherapy* 2004;38(9):1369-76.
109. Barnett N, Athwal D, Karen R. Medicines-related admissions: you can identify patients to stop that happening. *The Pharmaceutical Journal* 2011;286:471.
110. Hepler CD, Strand LM. Opportunities and responsibilities in pharmaceutical care. *Am J hosp pharm* 1990;47(3):533-43.
111. Schumock GT, Thornton JP. Focusing on the preventability of adverse drug reactions. *Hospital pharmacy* 1992;27(6):538-38.
112. Benkirane R, Soulaymani-Bencheikh R, Khattabi A, et al. Assessment of a new instrument for detecting preventable adverse drug reactions. *Drug safety* 2015;38(4):383-93.
113. Franklin BD, Birch S, Savage I, et al. Methodological variability in detecting prescribing errors and consequences for the evaluation of interventions. *Pharmacoepidemiology and drug safety* 2009;18(11):992-99.
114. Basger BJ, Moles RJ, Chen TF. Development of an aggregated system for classifying causes of drug-related problems. *Annals of Pharmacotherapy* 2015;49(4):405-18.
115. Cipolle RJ, Strand LM, Morley PC. Pharmaceutical care practice: McGraw-Hill 1998.
116. WESTERLUND T, ALMARSDÓTTIR AB, MELANDER A. Drug-related problems and pharmacy interventions in community practice. *International Journal of Pharmacy Practice* 1999;7(1):40-50.
117. Williams M, Peterson GM, Tenni PC, et al. DOCUMENT: a system for classifying drug-related problems in community pharmacy. *International journal of clinical pharmacy* 2012;34(1):43-52.

118. Ruths S, Viktil K, Blix H. Classification of drug-related problems. *Tidsskrift for den Norske laegeforening: tidsskrift for praktisk medicin, ny raekke* 2007;127(23):3073-76.
119. Crisp GD, Burkhart JI, Esserman DA, et al. Development and testing of a tool for assessing and resolving medication-related problems in older adults in an ambulatory care setting: the individualized medication assessment and planning (iMAP) tool. *The American journal of geriatric pharmacotherapy* 2011;9(6):451-60.
120. PROGRESS partnership. PROGRESS Summer School in Prognosis Research - Concepts, Methods and Clinical Application: Keele Univeristy, UK, 2015.
121. Michels WM, Grootendorst DC, Verduijn M, et al. Performance of the Cockcroft-Gault, MDRD, and new CKD-EPI formulas in relation to GFR, age, and body size. *Clinical Journal of the American Society of Nephrology* 2010;5(6):1003-09.
122. The Renal Drug Database: CRC Press.
123. Traynor J, Mactier R, Geddes CC, et al. How to measure renal function in clinical practice. *BMJ: British Medical Journal* 2006;333(7571):733.
124. National Institute for Health and Care Excellence. Chronic kidney disease: early identification and management of chronic kidney disease in adults in primary and secondary care, Clinical guideline [CG73], 2008.
125. The Renal Association. About eGFR [Available from: https://renal.org/information-resources/the-uk-eckd-guide/about-egfr/ accessed November 2017.
126. Giannini EG, Testa R, Savarino V. Liver enzyme alteration: a guide for clinicians. *Canadian medical association journal* 2005;172(3):367-79.
127. Sarfati D. How Do We Measure Comorbidity? Cancer and Chronic Conditions: Springer 2016:35-70.
128. Salvi F, Miller MD, Grilli A, et al. A manual of guidelines to score the modified cumulative illness rating scale and its validation in acute hospitalized elderly patients. *Journal of the American Geriatrics Society* 2008;56(10):1926-31.
129. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of chronic diseases* 1987;40(5):373-83.
130. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. *Medical care* 1998;36(1):8-27.
131. World Health Organization. International Statistical Classification of Diseases and Related Health Problems 10th Revision Instruction manual 2010 [Available from: http://www.who.int/classifications/icd/en/.

132. Oxford English Dictionary: Oxford University Press, 2017.
133. National Patient Safety Agency. Patient safety alert, Promoting safer use of injectable medicines (NPSA/2007/20), 2007.
134. National Institute for Health and Care Excellence. Venous thromboembolism: reducing the risk for patients in hospital, clinical guideline [CG92], 2015.
135. WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index, 2017.
136. NHS DATA MODEL AND DICTIONARY Version 3, 2017.
137. Vilela P. Acute stroke differential diagnosis: Stroke mimics. *European Journal of Radiology* 2017
138. Saedder EA, Brock B, Nielsen LP, et al. Identifying high-risk medication: a systematic literature review. *European journal of clinical pharmacology* 2014;70(6):637-45.
139. Boeker EB, Ram K, Klopotowska JE, et al. An individual patient data meta-analysis on factors associated with adverse drug events in surgical and non-surgical inpatients. *British journal of clinical pharmacology* 2015;79(4):548-57.
140. Thomas SK, McDowell SE, Hodson J, et al. Developing consensus on hospital prescribing indicators of potential harms amenable to decision support. *British journal of clinical pharmacology* 2013;76(5):797-809.
141. Institute for Safe Medication Practices. ISMP HIGH-ALERT MEDICATIONS [Available from: https://www.ismp.org/Tools/highAlertMedicationLists.asp accessed August 2017.
142. Geeson C, Wei L, Franklin BD. Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study. *BMJ Open* 2017;7(6) doi: 10.1136/bmjopen-2017-017509
143. Franklin BD, Reynolds M, Sadler S, et al. The effect of the electronic transmission of prescriptions on dispensing errors and prescription enhancements made in English community pharmacies: a naturalistic stepped wedge study. *BMJ quality & safety* 2014;23(8):629-38.
144. Institute for Healthcare Improvement. Medication Reconciliation to Prevent Adverse Drug Events [Available from: http://www.ihi.org/Topics/ADEsMedicationReconciliation/Pages/default.aspx accessed July 2017.

145. Sunquest Integrated Clinical Environment™ (Sunquest ICE®) [Available from: https://www.sunquestinfo.com/laboratory-orders-results/integrated-clinical-environment-for-pathology-networks-software/ accessed January 2018.
146. Armstrong RA. When to use the Bonferroni correction. *Ophthalmic and Physiological Optics* 2014;34(5):502-08.
147. Vandenbroucke JP, Von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007;4(10):e297.
148. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj* 2009;338:b2393.
149. GOV.UK. English indices of deprivation 2015, 2015.
150. Owen LJ, Keevil BG. Does bilirubin cause interference in Roche creatinine methods? *Clinical chemistry* 2007;53(2):370-71.
151. Pai SH, Cyr-Manthey M. Effects of hemolysis on chemistry tests. *Laboratory Medicine* 1991;22(6):408-10.
152. Puaar SJ, Franklin BD. Impact of an inpatient electronic prescribing system on prescribing error causation: a qualitative evaluation in an English hospital. *BMJ Qual Saf* 2017:bmjqs-2017-006631.
153. Brown CL, Mulcaster HL, Triffitt KL, et al. A systematic review of the types and causes of prescribing errors generated from using computerized provider order entry systems in primary and secondary care. *Journal of the American Medical Informatics Association* 2017;24(2):432-40. doi: 10.1093/jamia/ocw119
154. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare* 2016;9:211.
155. NHS Employers. Agenda for Change pay scales - Annual 2017 [Available from: http://www.nhsemployers.org/your-workforce/pay-and-reward/agenda-for-change/pay-scales/annual accessed October 2017.
156. Health Careers. Agenda for change - pay rates.
157. Randolph JJ. Online Kappa Calculator [Computer software], 2008.
158. Warrens MJ. Inequalities between multi-rater kappas. *Advances in data analysis and classification* 2010;4(4):271-86.
159. Brennan RL, Prediger DJ. Coefficient kappa: Some uses, misuses, and alternatives. *Educational and psychological measurement* 1981;41(3):687-99.
160. Randolph JJ. Free-Marginal Multirater Kappa (multirater K [free]): An Alternative to Fleiss' Fixed-Marginal Multirater Kappa. *Online submission* 2005

161. Bland M. An introduction to medical statistics: Oxford University Press (UK) 2015.
162. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37(5):360-63.
163. Dean BS, Barber ND. A validated, reliable method of scoring the severity of medication errors. *American Journal of Health-System Pharmacy* 1999;56(1):57-62.
164. McHugh ML. Interrater reliability: the kappa statistic. *Biochemia medica* 2012;22(3):276-82.
165. British National Formulary: National Institute for Health and Care Excellence, 2017.
166. Plakogiannis R, Cohen H. Optimal low-density lipoprotein cholesterol lowering—morning versus evening statin administration. *Annals of Pharmacotherapy* 2007;41(1):106-10.
167. Wallace A, Chinn D, Rubin G. Taking simvastatin in the morning compared with in the evening: randomised controlled trial. *Bmj* 2003;327(7418):788.
168. Rashed AN, Neubert A, Tomlin S, et al. Epidemiology and potential associated risk factors of drug-related problems in hospitalised children in the United Kingdom and Saudi Arabia. *European journal of clinical pharmacology* 2012;68(12):1657-66.
169. Hohl CM, Yu E, Hunte GS, et al. Clinical Decision Rules to Improve the Detection of Adverse Drug Events in Emergency Department Patients. *Academic Emergency Medicine* 2012;19(6):640-49. doi: 10.1111/j.1553-2712.2012.01379.x
170. Hoaglin DC, Iglewicz B, Tukey JW. Performance of some resistant rules for outlier labeling. *Journal of the American Statistical Association* 1986;81(396):991-99.
171. Osborne JW, Overbay A. The power of outliers (and why researchers should always check for them). *Practical assessment, research & evaluation* 2004;9(6):1-12.
172. Cox NJ. EXTREMES: Stata module to list extreme values of a variable 2017 [Available from: https://ideas.repec.org/c/boc/bocode/s430801.html accessed December 2017.
173. Williams R. Outliers: University of Notre Dame; 2016 [Available from: https://www3.nd.edu/~rwilliam/stats2/l24.pdf accessed December 2017.
174. UCLA. LESSON 3 LOGISTIC REGRESSION DIAGNOSTICS: UCLA: Statistical Consulting Group; [Available from:

https://stats.idre.ucla.edu/stata/webbooks/logistic/chapter3/lesson-3-logistic-regression-diagnostics-2/ accessed December 2017.

175. UCLA. LOGISTIC REGRESSION ANALYSIS | STATA ANNOTATED OUTPUT: UCLA: Statistical Consulting Group; [Available from: https://stats.idre.ucla.edu/stata/output/logistic-regression-analysis/ accessed December 2017.
176. StataCorp. STATA MULTIPLE-IMPUTATION, REFERENCE MANUAL, RELEASE 13: Stata Press; [Available from: https://www.stata.com/manuals13/mi.pdf accessed December 2017.
177. UCLA. MULTIPLE IMPUTATION IN STATA: UCLA: Statistical Consulting Group; [Available from: https://stats.idre.ucla.edu/stata/seminars/mi_in_stata_pt1_new/ accessed December 2017.
178. Koutoumanou E, Wade A. Introduction to Logistic Regression: Centre for Applied Statistics Courses, Great Ormond Street Institute of Child Health, 2015, Version 7.
179. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *International journal of epidemiology* 1999;28(5):964-74.
180. StataCorp. mfp — Multivariable fractional polynomial models [Available from: https://www.stata.com/manuals13/rmfp.pdf accessed December 2017.
181. Kirkwood BR, Sterne JA. Essential medical statistics: John Wiley & Sons 2010.
182. Koutoumanou E, Wade A, Lee S. Introduction to Dealing with Missing Data: Centre for Applied Statistics Courses, Great Ormond Street Institute of Child Health, 2017, Version 5.
183. Groeneveld RA, Meeden G. Measuring skewness and kurtosis. *The Statistician* 1984:391-99.
184. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative dentistry & endodontics* 2013;38(1):52-54.
185. Cousens S. Analysis of Correlated Outcome Data, Advanced Course in Epidemiological Analysis: London School of Hygiene and Tropical Medicine, 2017.
186. StataCorp. Working with categorical data and factor variables: Stata Press; [Available from: https://www.stata.com/manuals13/u25.pdf accessed January 2018.

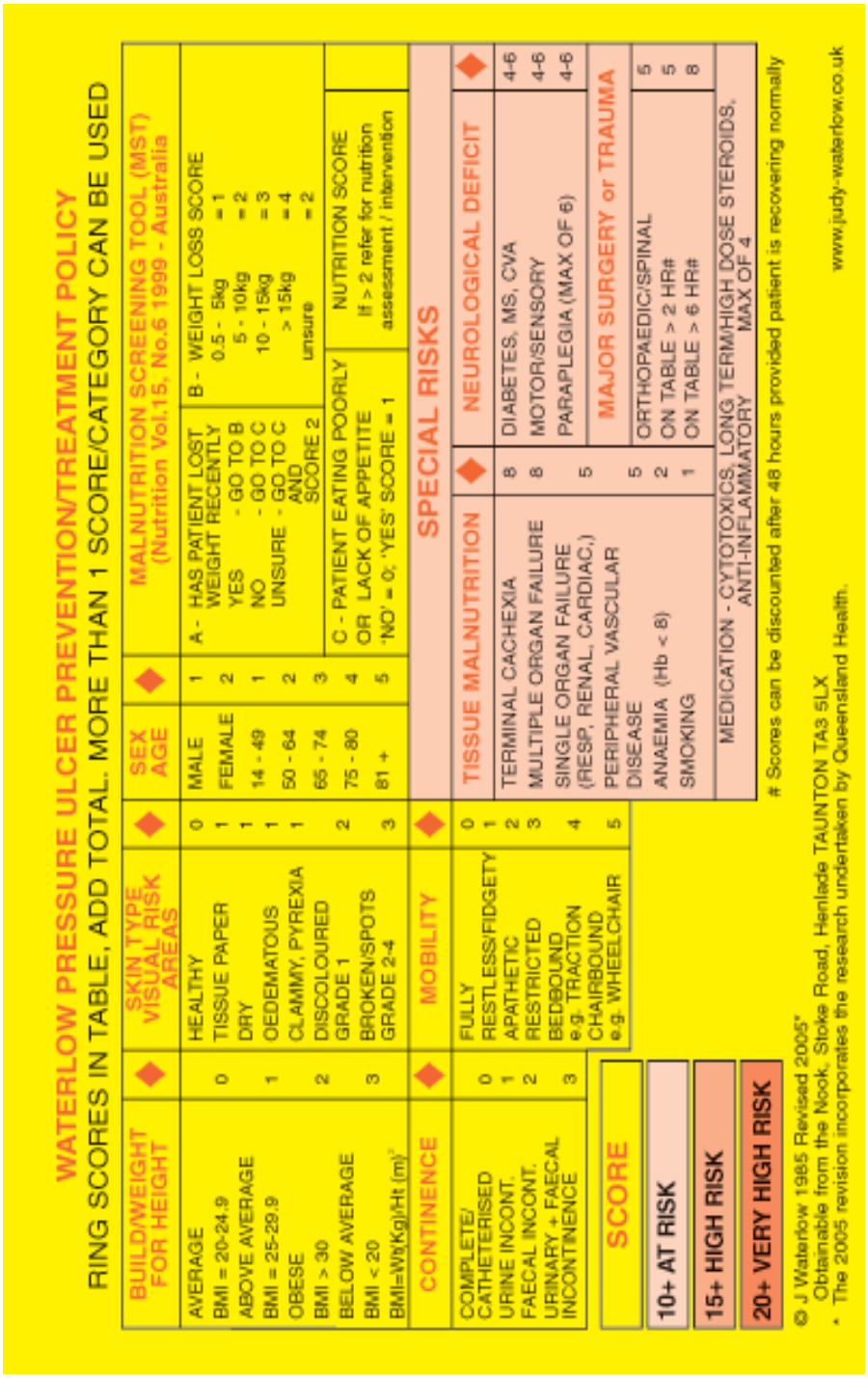
187. StataCorp. xtlogit — Fixed-effects, random-effects, and population-averaged logit models [Available from: https://www.stata.com/manuals13/xtxtlogit.pdf accessed January 2018.
188. StataCorp. quadchk — Check sensitivity of quadrature approximation [Available from: https://www.stata.com/manuals13/xtquadchk.pdf accessed January 2018.
189. Sarkar SK, Midi H, Rana S. Detection of outliers and influential observations in binary logistic regression: An empirical study. *Journal of Applied Sciences* 2011;11:26-35.
190. StataCorp. logit postestimation — Postestimation tools for logit: Stata Press; [Available from: https://www.stata.com/manuals13/rlogitpostestimation.pdf accessed January 2018.
191. StataCorp. xtlogit postestimation — Postestimation tools for xtlogit: Stata Press; [Available from: https://www.stata.com/manuals13/xtxtlogitpostestimation.pdf accessed January 2018.
192. StataCorp. mi predict — Obtain multiple-imputation predictions: Stata Press; [Available from: https://www.stata.com/manuals13/mimipredict.pdf accessed January 2018.
193. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. *Epidemiology (Cambridge, Mass)* 2010;21(1):128.
194. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European heart journal* 2014;35(29):1925-31.
195. Steyerberg E, Eijkemans M, Habbema J. Application of shrinkage techniques in logistic regression analysis: a case study. *Statistica Neerlandica* 2001;55(1):76-88.
196. Steyerberg EW, Eijkemans MJ, Harrell Jr FE, et al. Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Medical Decision Making* 2001;21(1):45-56.
197. Janssen K, Moons K, Kalkman C, et al. Updating methods improved the performance of a clinical prediction model in new patients. *Journal of clinical epidemiology* 2008;61(1):76-86.
198. StataCorp. stepwise — Stepwise estimation: Stata Pressstepwise — Stepwise estimation; [Available from: https://www.stata.com/manuals13/rstepwise.pdf accessed January 2018.

199. Garfield S, Jani Y, Jheeta S, et al. Impact of electronic prescribing on patient safety in hospitals: implications for the UK. *Stroke* 2018;13:57.
200. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Medical Decision Making* 2006;26(6):565-74.
201. Steyerberg EW, Vickers AJ. Decision curve analysis: a discussion. *Medical Decision Making* 2008;28(1):146-49.
202. Schmidt RL, Straseski JA, Raphael KL, et al. A risk assessment of the Jaffe vs enzymatic method for creatinine measurement in an outpatient population. *PloS one* 2015;10(11):e0143205.
203. gov.scot. Scottish Index of Multiple Deprivation, 2016.
204. Marchenko YV, Eddings W. A note on how to perform multiple-imputation diagnostics in Stata. *College Station, TX: StataCorp* 2011
205. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *The Annals of Family Medicine* 2004;2(3):204-08.
206. Altman DG, Vergouwe Y, Royston P, et al. Prognosis and prognostic research: validating a prognostic model. *Bmj* 2009;338:b605.
207. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Annals of emergency medicine* 1999;33(4):437-47.
208. Jones J, Hunter D. Consensus methods for medical and health services research. *BMJ: British Medical Journal* 1995;311(7001):376.
209. Specialist Pharmacy Service Medicines Use and Safety. Improving the Quality of Medicines Reconciliation, A Best Practice Resource and Toolkit, version 1.1 2017 [Available from: https://www.sps.nhs.uk/wp-content/uploads/2015/06/Medicines_Reconciliation_Best_Practice_Standards_Toolkit_Vs1.1_June-15-links-updated-Aug-17.pdf accessed March 2018.
210. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety: drawing together academic evidence and practical experience to produce a framework for safety measurement and monitoring: The Health Foundation 2013.
211. Rathert C, Brandt J, Williams ES. Putting the 'patient'in patient safety: a qualitative study of consumer experiences. *Health Expectations* 2012;15(3):327-36.
212. Lawton R, O'hara JK, Sheard L, et al. Can staff and patient perspectives on hospital safety predict harm-free care? An analysis of staff and patient survey data and routinely collected outcomes. *BMJ Qual Saf* 2015;24(6):369-76.

213. Elliott RA, Camacho E, Campbell F, et al. PREVALENCE AND ECONOMIC BURDEN OF MEDICATION ERRORS IN THE NHS IN ENGLAND.
214. PREVALENCE AND ECONOMIC BURDEN OF MEDICATION ERRORS IN THE NHS IN ENGLAND: Policy Research Unit in Economic Evaluation of Health and Care Interventions; 2018 [Available from: http://www.eepru.org.uk/prevalence-and-economic-burden-of-medication-errors-in-the-nhs-in-england-2/ accessed March 2018.
215. Matthews-King A. NHS medication errors contribute to as many as 22,000 deaths a year, major report shows: Independent; 2018 [Available from: http://www.independent.co.uk/news/health/nhs-medication-errors-deaths-prescription-drugs-jeremy-hunt-york-university-health-a8224226.html accessed March 2018.
216. Smyth C. Drug mistakes killing up to 22,300 patients a year: The Times; 2018 [Available from: https://www.thetimes.co.uk/article/drug-mistakes-killing-up-to-24-000-patients-a-year-q9lftn9bz accessed March 2018.
217. Lewis G. Next Steps for Risk Stratification in the NHS: NHS England; 2015 [Available from: https://www.england.nhs.uk/wp-content/uploads/2015/01/nxt-steps-risk-strat-glewis.pdf accessed March 2018.
218. Moons KG, Altman DG, Vergouwe Y, et al. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *Bmj* 2009;338:b606.
219. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Annals of internal medicine* 1999;130(6):515-24.
220. Saedder EA, Brock B, Nielsen LP, et al. Classification of drugs with different risk profiles. *Dan Med J* 2015;62:A5118.
221. Mullan N, Jennings A. Pharmacists' Use and Views of the Electronic Prescribing Web Portal [abstract]. GHP/UKCPA Joint National Conference 2013, 2014.
222. Safadeh M, Pazik L, Kavangh R. A baseline assessment of the pharmaceutical needs of adult patients admitted to Stoke Mandeville Hospital. *Clin Pharm* 2012
223. Intensive Care Society. Levels of Critical Care for Adult Patients, 2009.
224. The Shelford Group. Safer Nursing Care Tool implementation resouce pack.
225. Cousins DH, Gerrett D, Warner B. A review of medication incidents reported to the National Reporting and Learning System in England and Wales over 6 years (2005–2010). *British journal of clinical pharmacology* 2012;74(4):597-604.

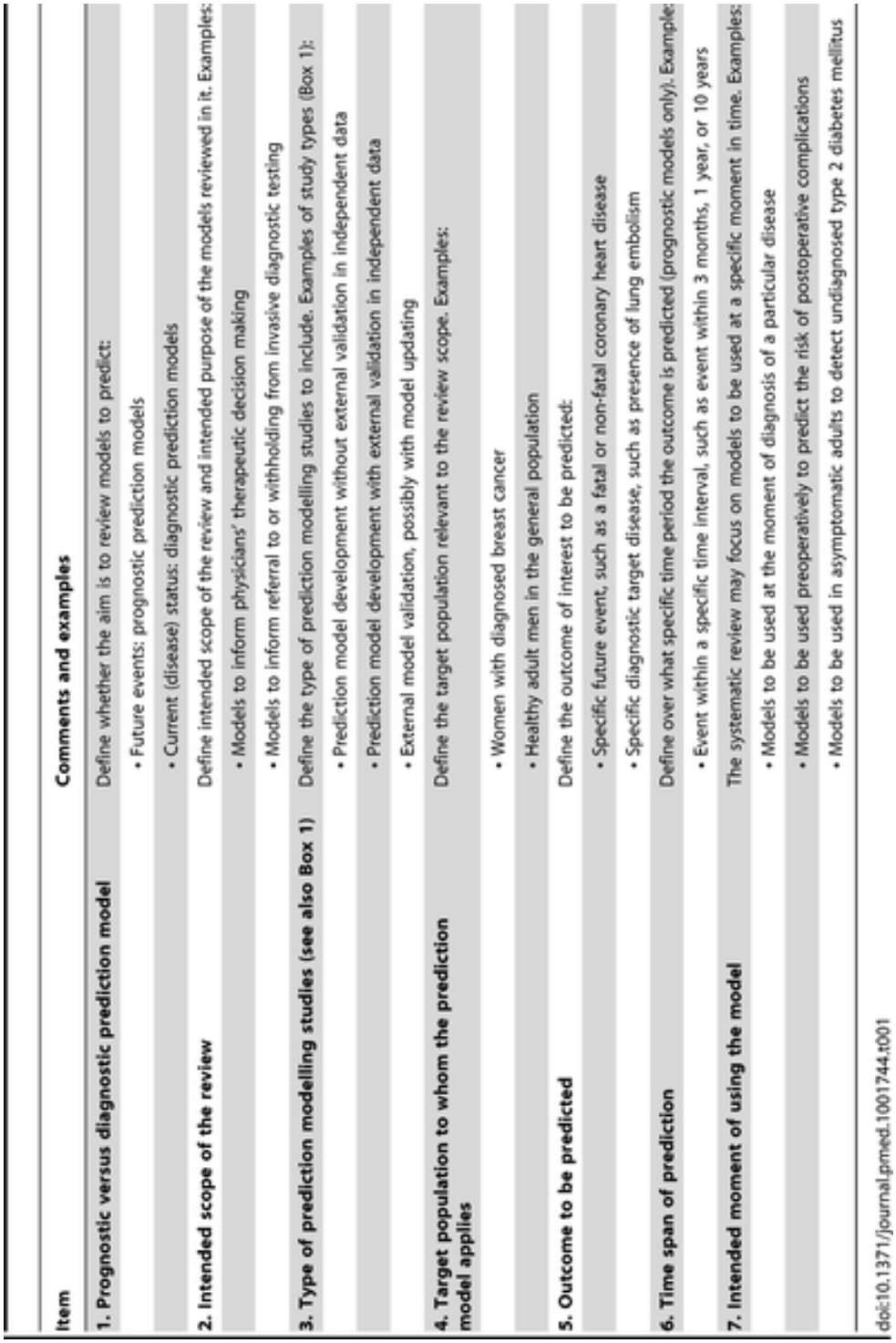
Appendices

Appendix A2.1: Waterlow score



Waterlow J. Waterlow Pressure Ulcer Prevention/Treatment Policy 2005 [Available from: http://www.judy-waterlow.co.uk/downloads/Waterlow%20Score%20Card-front.pdf]

Appendix A3.1: CHARMS guidance on key items to guide the framing of the review aim, search strategy, and study inclusion and exclusion criteria



Item	Comments and examples
1. Prognostic versus diagnostic prediction model	Define whether the aim is to review models to predict:
	• Future events: prognostic prediction models
	• Current (disease) status: diagnostic prediction models
2. Intended scope of the review	Define intended scope of the review and intended purpose of the models reviewed in it. Examples:
	• Models to inform physicians' therapeutic decision making
	• Models to inform referral to or withholding from invasive diagnostic testing
3. Type of prediction modelling studies (see also Box 1)	Define the type of prediction modelling studies to include. Examples of study types (Box 1):
	• Prediction model development without external validation in independent data
	• Prediction model development with external validation in independent data
	• External model validation, possibly with model updating
4. Target population to whom the prediction model applies	Define the target population relevant to the review scope. Examples:
	• Women with diagnosed breast cancer
	• Healthy adult men in the general population
5. Outcome to be predicted	Define the outcome of interest to be predicted:
	• Specific future event, such as a fatal or non-fatal coronary heart disease
	• Specific diagnostic target disease, such as presence of lung embolism
6. Time span of prediction	Define over what specific time period the outcome is predicted (prognostic models only). Examples:
	• Event within a specific time interval, such as event within 3 months, 1 year, or 10 years
7. Intended moment of using the model	The systematic review may focus on models to be used at a specific moment in time. Examples:
	• Models to be used at the moment of diagnosis of a particular disease
	• Models to be used preoperatively to predict the risk of postoperative complications
	• Models to be used in asymptomatic adults to detect undiagnosed type 2 diabetes mellitus

doi:10.1371/journal.pmed.1001744.t001

Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, et al. (2014) Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLOS Medicine 11(10): e1001744. https://doi.org/10.1371/journal.pmed.1001744

Appendix A3.2: Search strategy for EMBASE

Search performed August 2015:

Adverse medication-related outcomes

1. EMBASE; "med* related problem*".ti,ab; 555 results
2. EMBASE; "drug related problem*".ti,ab; 1752 results
3. EMBASE; exp ADVERSE DRUG REACTION/; 353726 results
4. EMBASE; "med* harm".ti,ab; 78 results
5. EMBASE; 1 OR 2 OR 3 OR 4; 355474 results

Prognostic factors/prediction tools

6. EMBASE; "clinical dec* rule*".ti,ab; 804 results
7. EMBASE; "clinical dec* tool*".ti,ab; 114 results
8. EMBASE; "predict* tool*".ti,ab; 4094 results
9. EMBASE; "predict* rule*".ti,ab; 2134 results
10. EMBASE; "prognos* model*".ti,ab; 3650 results
11. EMBASE; "prognos* tool*".ti,ab; 3354 results
12. EMBASE; exp RISK FACTOR/; 693724 results
13. EMBASE; "risk model*".ti,ab; 4977 results
14. EMBASE; "risk tool*".ti,ab; 301 results
15. EMBASE; "prognos* factor".ti,ab; 40678 results
16. EMBASE; "predict* variable*".ti,ab; 9659 results
17. EMBASE; 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16; 755086 results
18. EMBASE; 5 AND 17; 15569 results

Hospital

19. EMBASE; hospital*.ti,ab; 1240321 results
20. EMBASE; 18 AND 19; 2033 results

Additional search performed November 2015 to include medication errors:

Medication error terms

1. EMBASE; exp MEDICATION ERROR/; 14068 results.
2. EMBASE; "prescrib* error*".ti,ab; 893 results.
3. EMBASE; "drug* error*".ti,ab; 461 results.

Prognostic factors/prediction tools

4. EMBASE; exp RISK FACTOR/; 714881 results.
5. EMBASE; "predict* model*".ti,ab; 24866 results.

6. EMBASE; "risk model*".ti,ab; 5253 results.
7. EMBASE; "predict* tool*".ti,ab; 4305 results.
8. EMBASE; "clinical dec* tool".ti,ab; 77 results.
9. EMBASE; "clinical dec* rule".ti,ab; 490 results.
10. EMBASE; "predict* rule*".ti,ab; 2195 results.
11. EMBASE; "prognos* model*".ti,ab; 3854 results.
12. EMBASE; "prognos* tool*".ti,ab; 3548 results.
13. EMBASE; "risk tool*".ti,ab; 324 results.
14. EMBASE; "prognos* factor*".ti,ab; 99002 results.
14. EMBASE; "predict* variable*".ti,ab; 9985 results.
15. EMBASE; 1 OR 2 OR 3; 14693 results.
16. EMBASE; 4 OR 5 OR 6 OR 7 OR 8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14; 848238 results.
17. EMBASE; 16 AND 17; 691 results.

Appendix A3.3: QUIPS risk of bias assessment tool

Table. **Summary of the Bias Domains, Prompting Items, and Ratings of the QUIPS Tool***

Variable	Bias Domains					
	1. Study Participation	2. Study Attrition	3. Prognostic Factor Measurement	4. Outcome Measurement	5. Study Confounding	6. Statistical Analysis and Reporting
Optimal study or characteristics of unbiased study	The study sample adequately represents the population of interest	The study data available (i.e., participants not lost to follow-up) adequately represent the study sample	The PF is measured in a similar way for all participants	The outcome of interest is measured in a similar way for all participants	Important potential confounding factors are appropriately accounted for	The statistical analysis is appropriate, and all primary outcomes are reported
Prompting items and considerations†	a. Adequate participation in the study by eligible persons	a. Adequate response rate for study participants	a. A clear definition or description of the PF is provided	a. A clear definition of the outcome is provided	a. All important confounders are measured	a. Sufficient presentation of data to assess the adequacy of the analytic strategy
	b. Description of the source population or population of interest	b. Description of attempts to collect information on participants who dropped out	b. Method of PF measurement is adequately valid and reliable	b. Method of outcome measurement used is adequately valid and reliable	b. Clear definitions of the important confounders measured are provided	b. Strategy for model building is appropriate and is based on a conceptual framework or model
	c. Description of the baseline study sample	c. Reasons for loss to follow-up are provided	c. Continuous variables are reported or appropriate cut points are used	c. The method and setting of outcome measurement is the same for all study participants	c. Measurement of all important confounders is adequately valid and reliable	c. The selected statistical model is adequate for the design of the study
	d. Adequate description of the sampling frame and recruitment	d. Adequate description of participants lost to follow-up	d. The method and setting of measurement of PF is the same for all study participants		d. The method and setting of confounding measurement are the same for all study participants	d. There is no selective reporting of results
	e. Adequate description of the period and place of recruitment	e. There are no important differences between participants who completed the study and those who did not	e. Adequate proportion of the study sample has complete data for the PF		e. Appropriate methods are used if imputation is used for missing confounder data	
	f. Adequate description of inclusion and exclusion criteria		f. Appropriate methods of imputation are used for missing PF data		f. Important potential confounders are accounted for in the study design	
					g. Important potential confounders are accounted for in the analysis	
Ratings‡						
High risk of bias	The relationship between the PF and outcome is very likely to be different for participants and eligible nonparticipants	The relationship between the PF and outcome is very likely to be different for completing and noncompleting participants	The measurement of the PF is very likely to be different for different levels of the outcome of interest	The measurement of the outcome is very likely to be different related to the baseline level of the PF	The observed effect of the PF on the outcome is very likely to be distorted by another factor related to PF and outcome	The reported results are very likely to be spurious or biased related to analysis or reporting
Moderate risk of bias	The relationship between the PF and outcome may be different for participants and eligible nonparticipants	The relationship between the PF and outcome may be different for completing and noncompleting participants	The measurement of the PF may be different for different levels of the outcome of interest	The measurement of the outcome may be different related to the baseline level of the PF	The observed effect of the PF on outcome may be distorted by another factor related to PF and outcome	The reported results may be spurious or biased related to analysis or reporting
Low risk of bias	The relationship between the PF and outcome is unlikely to be different for participants and eligible nonparticipants	The relationship between the PF and outcome is unlikely to be different for completing and noncompleting participants	The measurement of the PF is unlikely to be different for different levels of the outcome of interest	The measurement of the outcome is unlikely to be different related to the baseline level of the PF	The observed effect of the PF on outcome is unlikely to be distorted by another factor related to PF and outcome	The reported results are unlikely to be spurious or biased related to analysis or reporting

PF = prognostic factor; QUIPS = Quality In Prognosis Studies.
* The **Supplement** (available at www.annals.org) shows the full QUIPS tool.
† Prompting items are to guide the user's judgment about risk of bias for each domain and are taken together to inform the overall judgment of potential bias and facilitate consensus among reviewers for each of the 6 domains. Some items may not be relevant to the specific study or the review research question; modification/clarification of the prompting items for the specific review question is encouraged.
‡ Each domain is rated as high, moderate, or low risk of bias considering the prompting items.

Hayden JA, van der Windt DA, Cartwright JL, et al. ASsessing bias in studies of prognostic factors. *Annals of Internal Medicine* 2013;158(4):280-86. doi: 10.7326/0003-4819-158-4-201302190-00009*

** Permission to use this image obtained from Annals of Internal Medicine*

Appendix A3.4: CHARMS checklist for critical appraisal and data extraction for systematic reviews of prediction modelling studies

Domain	Key items	General	Applicability	Risk of bias
Source of data	• Source of data (e.g., cohort, case-control, randomised trial participants, or registry data)		X	X
Participants	• Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centres, setting, inclusion and exclusion criteria)	X	X	
	• Participant description	X	X	
	• Details of treatments received, if relevant		X	X
	• Study dates	X	X	
Outcome(s) to be predicted	• Definition and method for measurement of outcome		X	X
	• Was the same outcome definition (and method for measurement) used in all patients?			X
	• Type of outcome (e.g., single or combined endpoints)	X	X	
	• Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?			X
	• Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?			X
	• Time of outcome occurrence or summary of duration of follow-up		X	
Candidate predictors (or index tests)	• Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease characteristics)	X		
	• Definition and method for measurement of candidate predictors		X	X
	• Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)		X	
	• Were predictors assessed blinded for outcome, and for each other (if relevant)?			X
	• Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)			X
Sample size	• Number of participants and number of outcomes/events	X		
	• Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable)			X
Missing data	• Number of participants with any missing value (include predictors and outcomes)	X		X
	• Number of participants with missing data for each predictor			X
	• Handling of missing data (e.g., complete-case analysis, imputation, or other methods)			X
Model development	• Modelling method (e.g., logistic, survival, neural networks, or machine learning techniques)	X		
	• Modelling assumptions satisfied			X
	• Method for selection of predictors for inclusion in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)			X
	• Method for selection of predictors during multivariable modelling (e.g., full model approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Information Criterion)			X
	• Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform shrinkage, penalized estimation)		X	X
Model performance	• Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals		X	
	• Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification improvement) and whether a priori cut points were used			X
Model evaluation	• Method used for testing model performance: development dataset only (random split of data, resampling methods, e.g., bootstrap or cross-validation, none) or separate external validation (e.g., temporal, geographical, different setting, different investigators)			X
	• In case of poor validation, whether model was adjusted or updated (e.g., intercept recalibrated, predictor effects adjusted, or new predictors added)		X	X
Results	• Final and other multivariable models (e.g., basic, extended, simplified) presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with standard errors or confidence intervals)	X	X	
	• Any alternative presentation of the final prediction models, e.g., sum score, nomogram, score chart, predictions for specific risk subgroups with performance	X	X	
	• Comparison of the distribution of predictors (including missing data) for development and validation datasets			X
Interpretation and Discussion	• Interpretation of presented models (confirmatory, i.e., model useful for practice versus exploratory, i.e., more research needed)	X	X	
	• Comparison with other studies, discussion of generalizability, strengths and limitations	X	X	

doi:10.1371/journal.pmed.1001744.t002

Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, et al. (2014) Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLOS Medicine 11(10): e1001744. https://doi.org/10.1371/journal.pmed.1001744

Appendix A3.5: Literature review – table of study characteristics

Source	Method of development / analysis	Country	Number of patients	Age of included patients (years)	Specialities included in study	Prospective?	Outcome event			Prediction tool developed?
							Type	Rated for severity?	Rated for preventability?	
Consensus studies										
Roten *et al.* (2010)[50]	Consensus	Switzerland	N/A	N/A	Internal medicine & geriatrics (for validation)	N/A	MRP	N/A	N/A	Yes
Cottrell *et al.* (2013)[48]	Consensus	Scotland	N/A	N/A	All inpatients	N/A	MRP	N/A	N/A	Yes
Falconer *et al.* (2014)[49]	Consensus	New Zealand	N/A	N/A	All inpatients	N/A	ADE & ME	N/A	N/A	Yes
Kaufmann *et al.* (2015)[38]	Consensus	Switzerland	N/A	N/A	All inpatients	N/A	MRP	N/A	N/A	No
Saedder *et al.* (2016)[46]	Consensus	Denmark	N/A	N/A	Internal medicines & orthopaedics	N/A	ME	N/A	N/A	Yes
Hickson *et al.* (2016)[51]	Consensus	England	N/A	N/A	All inpatients	N/A	MRP	N/A	N/A	Yes
Statistical studies										
Onder *et al.* (2010)[41]	Logistic regression	Italy	5,936	Over 65	Elderly	No	ADR	No	No	Yes

Continued from previous page…

Source	Method of development / analysis	Country	Number of patients	Age of included patients (years)	Specialities included in study	Prospective?	Outcome event			Prediction tool developed?
							Type	Rated for severity?	Rated for preventability?	
Tangiisuran *et al.* (2014)[42]	Logistic regression	England	690	65 & over	Elderly care & stroke	Yes	ADR	No	No	Yes
Kiguba *et al.*(2017)[43]	Logistic regression	Uganda	762	18 & over	Medical & gynaecology	Yes	ADR	No	No	Yes
McElnay *et al.* (1997)[44]	Logistic regression	Ireland	929	Over 65	Medical, surgical, cardiology & geriatrics	Yes	ADE	No	No	Yes
Trivalle *et al.* (2011)[45]	Logistic regression	France	526	Over 80	Geriatric rehabilitation	Yes	ADE	No	No	Yes
Nguyen *et al.* (2017)[47]	Logistic regression	France	1,408	Over 17	Medical & surgical	Yes	ME	Yes	No	Yes
Urbina *et al.* (2014)[52]	Logistic regression	Spain	8,713	19 & over	Medical, surgical & maternity	Yes	MRP	No	No	Yes
Bates *et al.* (1999)[76]	Logistic regression	USA	2,759	Not stated	Medical, ITU & surgical	Yes	ADE	Yes	Yes	No
Van den Bemt *et al.* (2000)[77]	Logistic regression	Nether-lands	538	19-97	Internal medicine	Yes	ADE	No	No	No
Blix *et al.* (2004)[78]	Log-linear regression	Norway	827	15-98	Internal medicine & rheumatol-ogy	Yes	MRP	No	No	No

Continued from previous page…

Source	Method of development / analysis	Country	Number of patients	Age of included patients (years)	Specialities included in study	Prospective?	Outcome event			Prediction tool developed?
							Type	Rated for severity?	Rated for preventability?	
Evans *et al.* (2005)[79]	Logistic regression	USA	68,835	Not stated	All inpatients	No	ADE	Yes	No	No
Johnston *et al.* (2006)[80]	Logistic regression	Canada	60,206	Not stated	All inpatients	No	ADE	No	No	No
Zopf *et al* (2008)[81]	Logistic regression	Germany	907	16-94	Internal medicine	Yes	ADR	No	No	No
Davies *et al.* (2009)[82]	Cox regression	England	3,695	Not stated	Medical & surgical	Yes	ADR	No	No	No
Zaal *et al.* (2010)[83]	Logistic regression	Nether-lands	592	Not stated	Internal medicine	Yes	ME*	Yes	No	No
Munoz-Torrero *et al.* (2010)[84]	Logistic regression	Spain	405	15-102	Internal medicine	Yes	ADR	No	No	No
Dequito *et al.* (2011)[85]	Logistic regression	Nether-lands	603	Not stated	Geriatrics & internal medicine	Yes	ADE	No	Yes	No
Ben-Yehuda *et al.* (2011)[86]	Logistic regression	Israel	274	53-84	Internal medicine	No	ME*	Yes	No	No

Continued from previous page…

Source	Method of development / analysis	Country	Number of patients	Age of included patients (years)	Specialities included in study	Prospective?	Outcome event			Prediction tool developed?
							Type	Rated for severity?	Rated for preventability?	
Beckett *et al.* (2012)[87]	Logistic regression	USA	340	Not stated	All inpatients	No	ADE	Yes	No	No
O'Connor *et al.* (2012)[88]	Logistic regression	Ireland	513	Over 65	General medical & surgical	Yes	ADR	No	No	No
Sikdar *et al.* (2012)[89]	Logistic regression	Canada	64,446	Over 65	All inpatients	No	ADR	No	No	No
Wilmer *et al.* (2015)[90]	Logistic regression	Nether-lands	131	Not stated	Pulmonary, cardiology & rehabilitation	Yes	MRP	No	No	No
Ashcroft *et al.* (2015)[91]	Logistic regression	England	26,019	Not stated	All inpatients	Yes	ME†	Yes	No	No
Ayalew *et al.* (2015)[72]	Logistic regression	Ethiopia	225	Not stated	Medical	Yes	MRP	No	No	No
Lenssen *et al.* (2016)[73]	Poisson regression	Germany	306	18-97	Urology, neurology & gastroenterol-ogy	Yes	MRP	No	No	No

* Prescribing and transcribing errors only
† Prescribing errors only

N/A = not applicable, ADR = adverse drug reaction, ADE = adverse drug event, ME = medication error, MRP = medication related problem, ITU = intensive therapy unit

Appendix A3.6: Risk of bias assessment for prognostic factor studies using QUIPS tool

Source	Domains	Ratings
Bates et al. (1999)[76]	Study participation	**Moderate** (Intensive care units intentionally oversampled which may skew data; no explanation of how the 2 study sites selected)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-points / categorisation used for laboratory results; proportion of patients with missing data not reported)
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	Low
Van den Bemt et al. (2000)[77]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards, or reasons for selecting the 2 study sites)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Age and number of medicines categorised during data collection; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (Outcome data collected from spontaneous reports from doctors & nurses, & from interviews with patients - therefore may not be comprehensive (in terms of reliability of spontaneous reporting, & ability of patients to correctly identify ADEs). This is shown by limited overlap (11%) among all 3 sources; no information given regarding the number of patients unable to be interviewed due to communication difficulties)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	Low
Blix et al. (2004)[78]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (A number of prognostic factors were considered together as 'clinical / pharmacological' factors therefore the impact of individual factors cannot be assessed; timing of data collection in relation to admission not stated; cut-points used for quantitative data; proportion of patients with missing data not reported)
	Outcome measurement	Low
	Study	N/A (exploratory study)

	confounding	
	Statistical analysis & reporting	**Moderate** (Limited data presented to assess the adequacy of the analytic strategy)
Evans et al. (2005)[79]	Study participation	Low
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-points used for renal function; timing of data collection for 'concomitant drugs' unclear in relation to admission; 16% of data missing – imputed using multiple regression imputation)
	Outcome measurement	**Moderate** (Retrospective study over a 10-year period therefore data not collected for study purpose which may affect the reliability of outcome data (i.e. only pharmacist verified ADEs included in the analysis, but no information provided on the number of ADE alerts generated but not pharmacist verified)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Continuous variables only categorised if linearity not supported – but unclear how categories selected)
Johnston et al. (2006)[80]	Study participation	Low
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Categorisation used for age; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (Data collected using voluntary reporting & analysed retrospectively, which may limit reliability of data collection; reports excluded if patient unidentifiable, or insufficient details provided – no information provided on the number excluded for these reasons)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Limited data presented to assess the adequacy of the analytic strategy; 93 prognostic factors identified for initial analysis & recursive partitioning used to reduce number prior to conducting regression analysis; authors acknowledge that 'drugs not included in the study may have proven to be good, if not better, indicators of the outcome')
Zopf et al (2008)[81]	Study participation	**Moderate** (Relatively small sample; no explanation of how the 2 study sites selected; no description of dates / duration of data collection)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Prognostic factor data collected at admission therefore does not account for changes during the admission; unclear definition for 'number of drugs during hospital stay'; proportion of patients with missing data not reported)
	Outcome measurement	Low

	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the multivariable regression based on univariable analysis)
Davies et al. (2009)[82]	Study participation	Low
	Study attrition	Low
	Prognostic factor measurement	Low
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the Cox regression based on univariable analysis; multivariable analysis carried out for 10% of patients (i.e. 374))
Zaal et al. (2010)[83]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-point used for renal function; unclear definition for 'number of medication orders during hospital stay'; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (Only prescribing and transcribing errors recorded as outcomes for this study; inappropriate drug choices not actively assessed & only included were obvious)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Study compared medication errors that 'cause harm' with those that 'do not cause harm', but only limited conclusions could be drawn due to the limited power for errors leading to harm; prognostic factors selected for the multivariable regression based on univariable analysis)
Munoz-Torrero et al. (2010)[84]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards; no explanation of how the 2 study sites selected)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-points used for number of drugs, number of drugs with interaction, & serum albumin; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Cut-point used to assess impact of a change in renal function during hospitalisation (>20%); prognostic factors

		selected for the multivariable regression based on univariable analysis)
Dequito et al. (2011)[85]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards; no explanation of how the 2 study sites selected)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Timing of data collection in relation to admission not stated; unclear definition for 'number of medication orders'; proportion of patients with missing data not reported)
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the multivariable regression based on univariable analysis)
Ben-Yehuda et al. (2011)[86]	Study participation	**Moderate** (Relatively small sample with cases matched 1:1 with controls; no description of included versus excluded wards)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Data on 'number of medications' collected at admission therefore does not account for changes during the admission; comorbidity score & number of medications categorised; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (Outcomes identified during twice weekly routine review of patients' charts & nurses' notes by investigators. This could result in missing data e.g. due to patient discharges, events that occur between scheduled review visits, or poor documentation in nursing records)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the multivariable regression based on univariable analysis)
Beckett et al. (2012)[87]	Study participation	**Moderate** (Relatively small sample with cases matched 1:1 with controls; controls randomly selected (i.e. matched based on severity of medication error only); no description of included versus excluded hospitals)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-points used for quantitative data; timing of data collection in relation to admission variable; proportion of patients with missing data not reported)
	Outcome measurement	**High** (Outcome data collected from voluntary reporting system)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the multivariable regression based on univariable analysis; only factors with 'sufficient entries in the data categories' were included in the analysis)

O'Connor et al. (2012)[88]	Study participation	**Moderate** (Relatively small sample)
	Study attrition	Low
	Prognostic factor measurement	Low
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Prognostic factors selected for the multivariable regression based on univariable analysis; age, renal function & liver disease categorised for analysis)
Sikdar et al. (2012)[89]	Study participation	Low
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Data on the 'number of drugs' not available for analysis as drug data not recorded in hospital database, therefore only descriptive analysis of drugs causing outcome available; comorbidity score categorised)
	Outcome measurement	**High** (Outcomes obtained from hospital coding data (i.e. external causes of injury associated with the use of drug) therefore reliant on the quality of coding; study focussed on the most severe outcomes, which represent a small proportion of overall ADRs)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Categorisation used for age ranges)
Wilmer et al. (2015)[90]	Study participation	**High** (Very small sample (131 patients); no description of included versus excluded wards; no inclusion / exclusion criteria given)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Cut-points used for quantitative data; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**High** (Method to identify outcomes not stated; outcome assessments performed by different physicians / pharmacists due to organisational limitations)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Limited data presented to assess the adequacy of the analytic strategy)
Ashcroft et al. (2015)[91]	Study participation	Low
	Study attrition	N/A (patient follow-up not required)
	Prognostic factor	**Moderate** (Proportion of patients with missing data not reported)

	measurement	
	Outcome measurement	Low
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Limited data presented to assess the adequacy of the analytic strategy)
Ayalew *et al.* (2015)[72]	Study participation	**Moderate** (Small sample; no details of number/details of patients who refused to participate; patients excluded if length of stay under 48 hours)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Limited data on definitions used; number of medicines categorised during data collection; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (MRPs identified remotely by researchers using standard reference databases i.e. not identified by routine care team; no description of clinical relevance of MRPs e.g. drug interactions accounted for 48% of all MRPs – unclear if these were restricted to those that are clinically relevant)
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	**Moderate** (Limited data presented to assess the adequacy of the analytic strategy)
Lenssen et al. (2016)[73]	Study participation	**Moderate** (Relatively small sample; no description of included versus excluded wards; unclear if all consecutive admissions included)
	Study attrition	Low
	Prognostic factor measurement	**Moderate** (Unclear definition used for 'number of drugs'; timing of data collection in relation to admission not stated; proportion of patients with missing data not reported)
	Outcome measurement	**Moderate** (Outcomes only recorded if action taken by healthcare team to implement advice (i.e. 77% of recommendations))
	Study confounding	N/A (exploratory study)
	Statistical analysis & reporting	Low

N/A = not applicable

Appendix A3.7: Risk of bias assessment for prognostic model studies using CHARMS tool

The CHARMS assessment for each prognostic study is shown below:

Domain	Key items	Onder *et al.*[41]
DATA SOURCE	Source of data	Retrospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Patients receiving at least 1 medication during hospital stay. Patients receiving anticancer medication excluded (total 61). Written consent required (no details given on patients who refused to consent).
	Participant description	65 years and over; community & university-based hospitals throughout Italy (part of the GIFA database).
	Study dates	September-October 1993, 1995 & 1997.
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	Outcome & method for measurement clearly defined. Definite & probable ADRs (using Naranjo algorithm, score 5-12 points). All wards visited daily by study physicians, who spoke to nurses & doctors, & examined medical & nursing records. Verification of ADEs not discussed.
	Same outcome definition (and measurement) used in all patients?	Definition the same, but details of inter-rater reliability among study physicians not stated. No further consensus validation.
	Type of outcome	Single endpoint (ADR).
	Blinding used?	Assume not, as all data (outcomes and predictors) collected by study physicians.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Data collected on medication use (but individual drug classes not included in the modelling), diagnoses, anaemia, renal function, nutritional status, activity of daily living disability, falls, history of ADRs, number of comorbid conditions, heart failure, & liver disease.
	Definition and method for measurement of candidate predictors	Unclear how nutritional status scored (assessed by measuring body mass index & serum albumin). No details of definition used for comorbid conditions, heart failure or liver disease. Details of inter-rater reliability among study physicians not stated. Drugs classified according to ATC system.
	Timing of predictor measurement	Not stated
	Blinding used?	Assume not for development study as all data (outcomes and predictors) collected by study physicians. Validation study: all data collected by study physicians, but 'physicians not informed regarding variables entered in risk score'.

	Handling of predictors in the modelling	Categorisation used for number of medicines, with '≤5', & '5-7' used (i.e. overlap between categories). Dichotomisation used for all other variables (anaemia, renal failure, activity of daily living disability, falls, number of comorbid conditions, heart failure, liver disease, & previous ADRs).
SAMPLE SIZE	Number of participants & outcomes / events	Development study: 5936 participants; 383 outcomes (6.5%).
	Events Per Variable (EPV)	Unable to determine as the number of candidate predictors, indicator variables, transformations, and interaction terms not stated. EPV appears to be >10.
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data), but no details regarding potential multicollinearity. Not stated if continuous variables checked for linearity.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used; variables added & retained in model if P≤0.1
	Method for selection of predictors **during modelling**	Stepwise logistic regression (not stated if forward or backwards selection), predictors added & retained if P≤0.1
	Shrinkage used?	Not performed
MODEL PERFORM-ANCE	Calibration and Discrimination	Calibration not assessed. Discrimination reported as the area under the ROC curve. Area under ROC curve 0.71 (95% CI 0.68-0.73).
	Classification measures	Internal validation: sensitivity 68%, specificity 65% using a cut-point score between 3 & 4 (not selected *a priori*).
MODEL EVALUATION	Method used for testing model performance	Prospective validation study of 483 patients from 4 university hospitals in Europe, September-December 2008. Exclusion criteria <65 years & refused consent. Outcome frequency 11.6% (i.e. 56 outcomes). Area under ROC curve 0.70 (95% CI 0.63-0.78).
	In case of poor validation, model adjusted?	N/A
RESULTS	Final and other multivariable models	Odds ratios and confidence intervals given for variables retained in multivariable analysis. Regression equation not given.

	Any alternative presentation of the final prediction models?	Sum score developed, with risk score assigned to each variable based on odds ratio. Final risk score based on 6 factors: ≥ 4 comorbid conditions, heart failure, liver disease, number of drugs (≤5, 5-7, ≥8), previous ADR, & renal failure. Scoring varies with factor.
	Comparison of the distribution of predictors for development and validation datasets	Frequency / distributions not given, but in the discussion it states that the validation group were 'sicker' i.e. 'older, had more comorbid conditions, and had a higher rate of ADRs'.
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Further studies needed to validate tool in different populations & settings.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Results compared with other prognostic studies. Strength: first study to identify high-risk patients; size of database used; assessment of ADRs done by study personnel during hospital stay. Limitations: data on preventability not recorded; results not generalisable to younger adults, or persons living in the community; prescribing patterns & disease burden differ by country; data collected between 1993 & 1997.

ADR = adverse drug reaction, GIFA = Gruppo Italiano di Farmacoepidemiologia nell'Anziano (Italian Group of Pharmacoepidemiology in the Elderly), ROC = receiver-operating characteristic, ATC = Anatomical Therapeutic Chemical, CI = confidence interval, N/A = not applicable

Domain	Key items	Tangiisuran *et al.*[42]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	All patients 'systematically enrolled' therefore assume consecutive admissions. Excluded when self-poisoning suspected, patient transferred to another ward during weekend, admitted & discharged during weekend, died within 24 hours of admission, or medical notes not available for further investigation. Written consent required (no details on patients who refused to consent). 256 excluded from final analysis as younger than 65, 111 as discharged or died before end of study period (no mention of other exclusions).
	Participant description	Patients 65 years and over admitted to 4 care of the elderly wards (2 elderly care, 2 stroke).
	Study dates	January-March 2007 & January-March 2008.
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	Outcome & method for measurement clearly defined. Patients reviewed by primary investigator using trigger tool. Data collected from procedure & emergency department notes, physician progress notes, laboratory reports, medication records, nursing flow sheets, hospital discharge records, & multidisciplinary progress notes. Additionally, reports from healthcare providers & incident reports included for further review. All ADRs discussed with attending physician & hospital pharmacist to confirm interpretation. Causality determined by review by investigator & independent reviewer using Hallas algorithm (categorised as definite, probable, possible or unlikely. Likert scale also used to assess confidence of relationship. Events classified as definite, probable or possible, with >50% confidence rating were classified as ADR cases.
	Same outcome definition (and measurement) used in all patients?	Yes (definition same, all ADR data collected by principal investigator & verified with physician & pharmacist, & causality confirmed by consensus).
	Type of outcome	Single endpoint (ADR).
	Blinding used?	Assume not, as all data (outcomes and predictors) collected by primary researcher.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Data collected on demographics (age, gender, ethnic origin) diagnosis, details of previous admissions, length of stay, social setting (living alone, smoking, alcohol), disability (Barthel Index), cognitive status (Abbreviated Mental Test Score), biological markers, drug history, & comorbidities (using ICD-10 codes).

	Definition and method for measurement of candidate predictors	Definition only given for renal failure, disability, & cognitive status only.
	Timing of predictor measurement	Obtained within 48 hours of admission.
	Blinding used?	No (all data collected by researcher).
	Handling of predictors in the modelling	Dichotomisation used for all variables.
SAMPLE SIZE	Number of participants & outcomes / events	690 participants; 86 outcomes (12.5%).
	Events Per Variable (EPV)	Unable to determine as the number of candidate predictors, indicator variables, transformations, and interaction terms not stated, but appears to be significantly <10 EPV.
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data). Selected variables assessed for multicollinearity (none found). Not stated if continuous variables checked for linearity.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used, predictors included in model if P<0.05. In addition, variables identified as being important in other studies with P<0.25 included. Variables present in <5% of study population omitted from analysis.
	Method for selection of predictors **during modelling**	Backwards elimination with removal criteria set at P=0.1. Process then repeated with forward selection procedure.
	Shrinkage used?	Not performed
MODEL PERFORMANCE	Calibration and Discrimination	Hosmer-Lemeshow test to assess overall fit (found to be satisfactory) & effect size measured using Nagelkerke R^2 (low, indicating small effect size). Discrimination reported as the area under the ROC curve. Area under ROC curve 0.74 (95% CI 0.68-0.79).
	Classification measures	Internal validation: sensitivity 80%, specificity 55%. Cut-point score >1 (not selected *a priori*).

MODEL EVALUATION	Method used for testing model performance	Prospective validation study of 483 patients from 4 European centres (September 2008-December 2008). Exclusion criteria identical to development dataset, plus patients receiving chemotherapy. Naranjo algorithm rather than Hallas criteria. Only events classified as definite or probable (score ≥5) considered to be drug related. Pair assessment not conducted. Outcome frequency 11.6% (i.e. 56 outcomes). Sensitivity 84%, specificity 43% with cut-point score >1 Area under ROC curve 0.73 (95% CI 0.66-0.80).
	In case of poor validation, model adjusted?	N/A
RESULTS	Final and other multivariable models	Intercept, regression coefficients, odds ratios and confidence intervals given for variables retained in multivariable analysis.
	Any alternative presentation of the final prediction models?	Results compared with other prognostic studies. Strength: first study to identify high-risk patients; size of database used; assessment of ADRs done by study personnel during hospital stay. Limitations: data on preventability not recorded; results not generalisable to younger adults, or persons living in the community; prescribing patterns & disease burden differ by country; data collected between 1993 & 1997.
	Comparison of the distribution of predictors for development and validation datasets	Given for length of stay (12 versus 10 days), number of medicines (rather than ≥8 medicines), & hyperlipidaemia (12.2 versus 28%). Not given for white cell count.
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Model validated across continental Europe, showing it is a reasonable robust, & may be applied to patients from other geographical locations & perhaps different healthcare systems with similar demographics. Need to establish if clinical judgement alone is more accurate, also need to test usability, & if use of the model would improve safety, & whether they are any humanistic & cost implications.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Discrimination comparable to Onder, sensitivity better, specificity lower, but Onder's model only considers definite & probable ADRs, whereas Tangiisuran included possible ADRs. Previous research identified age, previous ADR, gender, heart failure, prior bleeding on admission, renal impairment, use of certain drug classes, abnormal laboratory results as prognostic, but these not retained in their model. Strengths: Model simple to use, acceptable goodness of fit & good discrimination, good sensitivity, validation performed. Limitations: low specificity, length of stay is *a posteriori* observation, different casualty tool used for development & validation, may not be applicable to countries outside of Europe, risk of information bias.

ICD = International Classification of Diseases, N/A = not applicable

Domain	Key items	Kiguba *et al.*[43]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Systematic random sampling procedure; 3 new admissions per day on long-stay wards, 6 per day on short stay wards. Eligibility criteria not stated. Written consent required (no details given on patients who refused to consent).
	Participant description	18 years & over; 3 medical & 1 gynaecology ward.
	Study dates	December 2013-April 2014.
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	Outcome & measurement clearly defined. Possible and probable ADRs (using Naranjo algorithm, although cut-scores not stated). Possible ADRs included possible, probable & definite ADRs, probable included probably & definite. Data collected daily by 4 research teams (comprising doctor, pharmacist & nurse, who received training prior to study). Data captured from clinical notes, clinical examination & patient interviews. Consensus agreement on ADR causality reached by committee including principal author.
	Same outcome definition (and measurement) used in all patients?	Partly (definition same, & ADRs verified with physician & pharmacist, & causality confirmed by consensus, but unclear whether inter-rater reliability checked among research teams).
	Type of outcome	Single endpoint (ADR).
	Blinding used?	Assume not, as all data (outcomes and predictors) collected by research teams.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	No details given on predictor data collected. Final models included age, gender, number of conventional medicines, Charlson's comorbidity index, preadmission herbal medicines use, HIV-positive serostatus, hospitalisation in previous 3 months & gynaecology ward. Data on drug use collected, and analysed to identify drug classes associated with ADRs, but do not appear to have been used in model development.
	Definition and method for measurement of candidate predictors	No definitions given.
	Timing of predictor measurement	Not stated
	Blinding used?	No (all data collected by research teams).

	Handling of predictors in the modelling	Cut-points used for number of conventional medicines and comorbidity score. Interaction assessed for HIV status and the 6 categorical variables - interaction between Charlson's index & HIV retained in model for possible ADRs.
SAMPLE SIZE	Number of participants & outcomes / events	762 participants; 194 possible ADRs (25.5%), 87 probable ADRs (11.4%).
	Events Per Variable (EPV)	Sample size calculated based on incidence of ADRs & 5% confidence interval. Unable to determine EPV as the number of candidate predictors not stated. If it is assumed that only the predictors included in the final model were used (i.e. 9 variables), EPV for possible ADRs = 21.5, EPV for probably ADRs = 9.7 (assuming no indicator variables, transformations, and interaction terms used).
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Low frequency missing data for binary categorical variables set to 'no'.
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data), but no details regarding potential multicollinearity. Limited information on checks for linearity. They state that 'results obtained for linearity were compared with those using categorisation by comparing log-likelihoods for regressors in logistic regression for natural logarithms of the odds on having experienced possible or probable ADR', but no results presented. Models stratified by HIV serostatus to assess for interaction with the 6 categorised predictors (significance level set at 1% level).
	Method for selection of predictors **for inclusion** in modelling	Not stated
	Method for selection of predictors **during modelling**	Not stated
	Shrinkage used?	Not performed

MODEL PERFORMANCE	Calibration and Discrimination	Not performed
	Classification measures	Not stated
MODEL EVALUATION	Method used for testing model performance	None
	In case of poor validation, model adjusted?	N/A
RESULTS	Final and other multivariable models	Two regression models developed (for probable ADRs, and possible ADRs). Intercept, odds ratios and confidence intervals given for variable included in multivariable analysis. Basic regression equation given for both models. Final models included age, gender, number of conventional medicines, Charlson's comorbidity index, preadmission herbal medicines use, HIV-positive serostatus, hospitalisation in previous 3 months & gynaecology ward (despite not all being statistically associated). Statistical significant predictors of probable ADRs = use of ≥ 6 conventional medicines & self-reported herbal medicine use. Statistical significant predictors of possible ADRs = use of ≥6 conventional medicines & self-reported herbal medicine use, hospitalisation in 3 months prior to admission, and gynaecology ward.
	Any alternative presentation of the final prediction models?	No
	Comparison of the distribution of predictors for development and validation datasets	N/A
INTERPRETATION AND DISCUSSION	Interpretation of presented models	More research needed to validate model.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Limitations: 'number of conventional medicines' may include medicines received after ADR occurred, suggest could have used Cox proportional hazards to track daily changes; clinical examination (rather than laboratory investigations) used as main method to identify ADRs.

HIV = human immunodeficiency virus, N/A = not applicable

Domain	Key items	McElnay *et al.*[44]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Consecutive admissions. Inclusion criteria: unplanned admission, written consent, receiving medication at time of admission. 72 patients did not meet eligibility criteria - but further details not given.
	Participant description	65 years and over; medical, surgical, cardiology & geriatric wards.
	Study dates	Not stated
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	Outcome described as definite & probable ADEs based on Naranjo algorithm (amended to include compliance assessment) with a score of 4 or more, although unclear how compliance assessed. Not clear who collected study data - but data collected using chart review, search of patient's computerised hospital records & structured patient interview. 50.2% of study group & 41.7% of validation group interviewed. Verification of ADEs not discussed.
	Same outcome definition (and measurement) used in all patients?	Partly (definition same, but unclear who collected the study data & whether ADEs verified, or inter-rater reliability checked).
	Type of outcome	Combined endpoint (ADE i.e. ADR plus compliance) - frequency of individual component not given.
	Blinding used?	Unclear who collected study data, therefore unable to assess.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Not clear who collected study data - but data collected using same methods as outcome assessment. Variables included demographic, patient history, laboratory tests, clinical examination & medication use. 47 predictors listed, plus medication (classified by pharmaceutical class), serum drug concentrations' & individual medical problems. Total number of variables investigated not stated.
	Definition and method for measurement of candidate predictors	No definitions given. Drugs classified according to pharmacological class.
	Timing of predictor measurement	Timing for 'number of medicines', physical examination & laboratory results not stated.
	Blinding used?	As with outcome, unclear who collected study data, therefore unable to assess.

	Handling of predictors in the modelling	Not stated, but potassium & glucose levels dichotomised in univariable results.
SAMPLE SIZE	Number of participants & outcomes / events	929 participants; number of outcomes not stated, but ADEs (all scores) recorded in 16% of patients, therefore ADEs with score of ≥4 lower than 148.
	Events Per Variable (EPV)	Unable to determine as the number of candidate predictors, indicator variables, transformations and interaction terms not stated, but as 22 variables had P≤0.05 (i.e. lower than value for model inclusion) the EPV <7
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated. Percentage of patients interviewed is given, but not which predictors collected by this method, or how missing data handled during analysis.
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data). Variables assessed for multicollinearity (pairwise interaction terms for the model evaluated with entry set at P=0.05) but results not reported. Not stated if continuous variables checked for linearity.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used, predictors included in model if P<0.25 & occurred in >5% of study population. In addition all drug classes included in modelling.
	Method for selection of predictors **during modelling**	Backwards elimination with removal criteria set at P=0.15 then P=0.05
	Shrinkage used?	Not performed
MODEL PERFORMANCE	Calibration and Discrimination	Not performed
	Classification measures	Not stated
MODEL EVALUATION	Method used for testing model performance	Prospective validation study of 204 patients (inclusion criteria as above). Number of outcomes not stated. Profile of the validation group was similar to that recorded for phase 1 study group. Study site & dates not stated. Sensitivity 40.5%, specificity 69%, overall accuracy 63% (when cut-point set to model-predicted probability of 0.3, selected to produce optimal sensitivity & specificity in validation group - i.e. not selected *a priori*).

	In case of poor validation, model adjusted?	No
RESULTS	Final and other multivariable models	Intercept, regression coefficients, odds ratios and confidence intervals given for variable retained in multivariable analysis. Final model based on 7 factors: antidepressants, digoxin, gastrointestinal problems, abnormal serum potassium, 'thinks drugs were responsible for hospitalisation', angina, & congestive obstructive pulmonary disease. Scoring system not stated.
	Any alternative presentation of the final prediction models?	No
	Comparison of the distribution of predictors for development and validation datasets	Not stated
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Sensitivity indicates model is not a satisfactory predictor of ADEs.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Not discussed

ADE = adverse drug event

Domain	Key items	Trivalle *et al.*[45]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Consecutive admissions. Patients excluded if they were in the intervention group of the original study (54), or not present during all 4 weeks of the study (17). Not stated if written consent required.
	Participant description	65 years and over; 16 geriatric centres.
	Study dates	Not stated
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	Outcome clearly defined, but standardised system (Naranjo or Hallas) not used. ADEs identified using a standardised checklist (32 items, e.g. sleepiness, fall, vomit, diarrhoea) by nurses & a weekly review of all charts by investigators. Nurses & residents asked to report incidents to investigators. Investigators visited each unit weekly. Instructions given on how to complete data collection forms. Investigators reviewed all charts weekly. Standardised method used by group of reviewers to classify ADEs (either not ADE or probable ADE).
	Same outcome definition (and measurement) used in all patients?	Partly (definition same, & consensus group used to verify ADEs, but not clear if inter-rater reliability checked among investigators).
	Type of outcome	Single endpoint (ADE).
	Blinding used?	Assume not, as data collected by nurses & investigators (who also reviewed patients' charts weekly).
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Specific details of predictors other than those identified by univariable analysis (cardiovascular disease, respiratory disease, recent VTE, antihypertensives, & antidepressants), or included in final model (number of medicines, antipsychotics, & recent anticoagulation) not given, although some mentioned in discussion (e.g. heart failure, liver disease, renal failure, number of diseases, number of comorbidities & age not associated with ADEs).
	Definition and method for measurement of candidate predictors	No information on which data collected. Medicines classified using ATC system, no other definitions given.
	Timing of predictor measurement	All risk factor data collected as of date of the ADE, unclear when data collected for patients who did not experience an ADE.
	Blinding used?	No details of who collected this data, but assume

		investigators. As data collected as of date of ADE assume that not blinded.
	Handling of predictors in the modelling	Categorisation used for number of medicines.
SAMPLE SIZE	Number of participants & outcomes / events	526 participants included an original study, & 223 outcomes identified (39%). For this study 54 patients (from original intervention group) excluded, plus 17 patients not present during all 4 weeks, therefore assume 455 participants. 152 outcomes (33.4%)
	Events Per Variable (EPV)	Number of predictors included in modelling not stated therefore unable to assess.
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data). Correlations between potential risk factors assessed, & highly correlated variables analysed in separate models. Linear trend checked for number of medications, no other details given.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used, predictors included in model if P<0.05 with a prevalence of at least 5%.
	Method for selection of predictors **during modelling**	Stepwise logistic regression (not stated if forward or backwards selection). Variables retained if P≤0.05
	Shrinkage used?	Not performed
MODEL PERFORMANCE	Calibration and Discrimination	Calibration not assessed. Discrimination reported as the area under the ROC curve. Area under ROC curve 0.70 (95% CI 0.65-0.74).
	Classification measures	Not stated
MODEL EVALUATION	Method used for testing model performance	The area under the ROC curves were 'calculated and bootstrapped', but no further details given.
	In case of poor validation, model adjusted?	N/A

RESULTS	Final and other multivariable models	Odds ratios and confidence intervals given for variables retained in multivariable analysis. Regression equation not given.
	Any alternative presentation of the final prediction models?	Sum scored developed with each item weighted in proportion to its regression coefficient. Final risk score based on 3 factors: number of drugs, antipsychotics, & recent anticoagulation. Scoring varies with factor. Risk groups not created, but risk of ADE estimated by bootstrapping. A score of <6 gives an ADE risk of 12% (95% CI 8-15%), score of 7-12 ≡ 28% (95% CI 19-36%), score 13-18 ≡ 35% (95% CI 28-43%), & >18 out of maximum 34 gives as estimated risk of ADE of 52% (95% CI 40-62%).
	Comparison of the distribution of predictors for development and validation datasets	N/A
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Future studies needed to validate score in other settings & countries.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Results compared with other prognostic studies, discrimination similar to Onder. Strengths: multicentre trial; standardised data collection; results analysed by multidisciplinary team. Limitations: sample size; results cannot be extrapolated to community; prescribing may be specific to country; geriatric units may have better prevention of ADEs, underestimating occurrence.

N/A = not applicable

Domain	Key items	Nguyen *et al.*[47]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Assume consecutive admissions (all patients included during study period). Eligibility criteria & number of excluded patients not stated. Consent not required.
	Participant description	17 years & over; 21 hospital units (4 surgical & 17 medical).
	Study dates	1-31 April 2014.
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	At least 1 clinically relevant ME during hospital stay (verified by a physician & pharmacist, who also classified clinical relevance using a European adaptation of NCC MERP index, disagreements resolved by third evaluator). Outcome & measurement clearly defined. Wards visited by 1 of 15 pharmacists, all trained in standardised medicine reconciliation, prescription analysis & ME reporting.
	Same outcome definition (and measurement) used in all patients?	Partly (definition same, & consensus group used to verify ADEs, but not clear if inter-rater reliability checked among pharmacist who collected data).
	Type of outcome	Single endpoint (medication error).
	Blinding used?	Assume not, as data collected by clinical pharmacists.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Data collected on age, sex, admission details, admission type (medical versus surgical), previous hospitalisation in 30 day, hour & day of admission, number of prescribed drugs, drug class, treatment initiated before admission, medication history available.
	Definition and method for measurement of candidate predictors	Admission details', 'number of prescribed drugs' & ' treatment initiated before entrance' not defined. Details of inter-rater reliability among pharmacists not stated. Drugs classified according to ATC system.
	Timing of predictor measurement	Not stated
	Blinding used?	No details of who collected this data, or if blinding used.
	Handling of predictors in the modelling	Continuous predictors (age & number of drugs) treated as continuous data in final model. Quadratic & cubic terms used for age (to account for non-linear relationship with outcome).
SAMPLE SIZE	Number of participants & outcomes / events	1408 participant; 365 outcomes (25.9%).

	Events Per Variable (EPV)	Unable to determine as the number of candidate predictors, indicator variables, transformations and interaction terms) not stated.
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, & observations independent (i.e. no paired data). No information on whether variables assessed for multicollinearity. Multiple fractional polynomials analysis used to take account of non-linearity of continuous variables.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used, predictors included in model if P<0.5
	Method for selection of predictors **during modelling**	Selection method not stated, but alpha risk threshold set at P<0.3
	Shrinkage used?	Bootstrapping used to correct optimistic performance.
MODEL PERFORMANCE	Calibration and Discrimination	Calibration plot presented, & discrimination reported as the c-statistic. Calibration: reported as 'good' (intercept equal to zero & slope equal to 1) but slight over-estimation of high probabilities; C-statistic 0.718 (95% CI: 0.689-0.748).
	Classification measures	Not stated
MODEL EVALUATION	Method used for testing model performance	Internal validation performed using bootstrapping. To evaluate the impact of the model in clinical practice they conducted a series of simulated randomised controlled trials to compare strategies for pharmaceutical intervention (using the model compared with age) using different levels of pharmacy coverage. In all coverage scenarios pharmacists more likely to identify MEs using model.
	In case of poor validation, model adjusted?	Intercept and regression coefficients corrected following bootstrapping. Corrected c-statistic 0.707, 95% CI not given.
RESULTS	Final and other multivariable models	Intercept, odds ratios and confidence intervals given for variables retained in multivariable analysis. Basic formula for regression equation given. Model includes 11 predictors.

		Medication error significantly associated with current treatment before admission, number of prescribed drugs & increasing age. Admission to surgical ward, admission within 30 days & psycholeptics almost achieved statistical significance. Other informative predictors also included in model (best possible medication history available, blood substitutes & perfusion solutions, admission from emergency room, admission time (day versus night), admission from outside institution).
	Any alternative presentation of the final prediction models?	No
	Comparison of the distribution of predictors for development and validation datasets	N/A
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Requires external validation & evaluation of concrete clinical outcomes.
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Discrimination comparable to Onder, Tangiisuran, Trivalle, & Urbina. Strengths: used TRIPOD guidelines, adjusted for optimism & assessed impact on practice; limitations: model does not consider biological markers, diagnostic categories or comorbidities.

ME = medication error, NCC MERP = National Coordinating Council for Medication Error Reporting and Prevention, N/A = not applicable

Domain	Key items	Urbina *et al.*[52]
DATA SOURCE	Source of data	Prospective cohort
PARTICIPANTS	Participant eligibility and recruitment method	Assume consecutive admissions. Patients excluded if 18 years or less (1128), admitted directly to critical care unit (733), or to emergency department without hospital admission / units without CPOE (1853). **NB numbers excluded include development & validation samples.** Consent not required.
	Participant description	Patients 18 years or less excluded; medical, surgical, & maternity wards.
	Study dates	January-August 2009.
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	MRPs detected through pharmacy warning system integrated in computerised medical history. When prescription issued the programme generated a series of alerts (causes of possible MRPs) based on drug information introduced & each patient's demographic characteristics & laboratory data. Alerts and prescriptions reviewed daily by team of pharmacists & detected MRPs that are considered clinically relevant & proposed intervention reported to prescriber. Data collected on causes of potential MRPs detected by pharmacy warning system & its classification (using PCNE classification system). Not clear whether all alerts used as outcomes, or just those verified by pharmacists. **NB prescription errors due to incorrect use of the CPOE system accounted for 23.9% of outcomes (highest of all categories of events).**
	Same outcome definition (and measurement) used in all patients?	Assume yes, is they are simply the alerts from the pharmacy warning system.
	Type of outcome	Combined endpoint (MRPs) - frequency of individual component given (using PCNE classification).
	Blinding used?	Appears to be automated alerts from system, therefore yes.
	Predictors part of outcome?	No
	Time of outcome occurrence	During admission
CANDIDATE PREDICTORS	Number and type of predictors	Data collected on demographics (age, sex), type of admission (urgent or elective), major diagnostic category, admitting department (surgical or medical), Charlson comorbidity index, diagnosis-related groups weight, obesity, cachexia, glomerular filtration rate, liver failure, number of distinct drugs received during admission, readmission related to prior admission, drugs classified according to ATC system.

	Definition and method for measurement of candidate predictors	No definition for liver failure, number of drugs received during admission, obesity, cachexia. Details of inter-rater reliability among study physicians not stated.
	Timing of predictor measurement	Not stated
	Blinding used?	No details of who collected these data, or if blinding used.
	Handling of predictors in the modelling	Charlson index, obesity, cachexia, renal failure & liver disease categorised.
SAMPLE SIZE	Number of participants & outcomes / events	Development study: 8713 admissions; 2425 outcomes (27.8%).
	Events Per Variable (EPV)	Not stated in text, but 34 variables listed in univariable results table giving 71 EPV (assuming no indicator variables, transformations and interaction terms used).
MISSING DATA	Number of participants with any missing value	Not stated
	Number of participants with missing data for each predictor	Not stated
	Handling of missing data	Not stated
MODEL DEVELOPMENT	Modelling method	Logistic regression
	Modelling assumptions satisfied	Binary dependent variable used, but not all observations independent (i.e. 8,713 admissions, 7,202 patients). No details regarding potential multicollinearity. Not stated if continuous variables checked for linearity.
	Method for selection of predictors **for inclusion** in modelling	Univariable analysis used, predictors included in model if P<0.1
	Method for selection of predictors **during modelling**	Variables whose exclusion did not significantly change model's verisimilitude or the coefficients of remaining variables were excluded.
	Shrinkage used?	Not performed
MODEL PERFORMANCE	Calibration and Discrimination	Hosmer-Lemeshow test to assess calibration. Discrimination reported as the area under the ROC curve. Results of Hosmer-Lemeshow test not stated. Area under ROC curve 0.778 (95% CI 0.768-0.789)
	Classification measures	Not stated

MODEL EVALUATION	Method used for testing model performance	Prospective validation study of 4058 consecutive admissions from same centre, September-December 2009. Outcome frequency 21.6% (i.e. 876 outcomes). Area under ROC curve 0.776 (95% CI 0.759-0.792). 'Model not significant' according to Hosmer Lemeshow's test (P=0.131).
	In case of poor validation, model adjusted?	N/A
RESULTS	Final and other multivariable models	Intercept, regression coefficients, odds ratios and confidence intervals given for variable retained in multivariable analysis. Basic regression equation given.
	Any alternative presentation of the final prediction models?	Sum score developed using variables retained in model, with risk score assigned to each variable based on odds ratio. Final risk score based on 14 factors: age >60 years, Charlson index =2, number of drugs >10, 6 'major diagnostic category (MDC) groups; 5 ATC groups. Scoring varies with factor. Age & number of medicines categorised using cut-points with highest sensitivity and specificity.
	Comparison of the distribution of predictors for development and validation datasets	Not stated
INTERPRETATION AND DISCUSSION	Interpretation of presented models	Not stated
	Comparison with other studies, discussion of generalizability, strengths and limitations.	Results compared with other prognostic studies. Strength: first study to design & validate a predictive score to detect MRPs in hospitalised patients; CPOE warning system includes information on drugs, diagnostic, & laboratory tests, allowing extensive & automatic patient monitoring & routine updating; large sample sizes. Limitations: comparison with other studies limited by definitions & outcomes used in other studies; results may not be reproducible to other hospitals as CPOE designed, & model validated, in staff at their centre.

MRP = medication related problem, PCNE = Pharmaceutical Care Network Europe, CPOE = computerised physician order system, N/A = not applicable

Appendix A3.8: Risk of bias assessment for consensus studies

Roten *et al.* (2010)[50]	**Summary of study:** Development of an electronic screening tool to identify patients at risk of MRPs.
	Study method: Six 'electronic queries' were formulated based on a literature review, clinical pharmacists' experience, a list if queries used at Brigham and Women's Hospital, Boston USA, and programming feasibility.
	Expert group: Clinical pharmacists at Hôpital du Valais. No further details given regarding number, experience of participants, or process used.
	Testing/validation: Prospective observational study of 501 patients to compare identification of patients using the electronic flags to a manual check by pharmacists. Sensitivity 85.1%, specificity 60.4%.
Cottrell *et al.* (2013)[48]	**Summary of study:** Development of an automated electronic screening tool using data from an electronic prescribing system to assist pharmacists target patients in greatest need.
	Study method: Initial priority criteria identified through discussions with clinical pharmacy team at University Hospital Ayr, and an initial scoring system developed to categorise patients as low, medium, or high-risk. This was refined during a 15-week evaluation period through weekly meetings of the pharmacist leading the project and a minimum of four clinical pharmacists. Feedback was also obtained from members of the clinical pharmacy team (who used the screening tool within their day-to-day work), resulting in *ad hoc* amendment (with contentious changes requiring the consensus of the clinical teams).
	Expert group: Iterative process involving the clinical pharmacy team members from medical and surgical specialities within University Hospital Ayr.
	Testing/validation: 'Testing phase' reported (no details given). Patients scored by algorithm as 'high-risk' compared to those identified by 'traditional ward round'. Match reported[71].
Falconer *et al.* (2014)[49]	**Summary of study:** Development of a software-based tool, the Assessment of Risk Tool (ART) to prioritise inpatients for ADE prevention.
	Study method: The system was designed with 38 risk 'flags, identified from the literature. Each flag was assigned a score by senior pharmacists at Middlemore Hospital using a group consensus process. Each patient's scores from all triggered flags are then summed to categorise patient as low, medium or high-risk for MEs and ADEs. Allocation of scores to risk groups was on the basis of staffing constraints, so that the top 10% (by ART score) would be categorised at high-risk, those in the next 15th percentile were medium-risk, and the remainder were low-risk.
	Expert group: Senior pharmacists at Middlemore Hospital. No further details given regarding number, experience of participants, or process used.
	Testing/validation: None reported
Kaufmann *et al.* (2015)[38]	**Summary of study:** Expert panel to gather risk factors for MRPs.
	Study method: Triangulation process using a mixed method approach to identify risk factors for MRPs. This included the nominal group technique (NGT), literature review, and a two-round Delphi survey.
	Expert group: Two hospital physicians, one emergency physician, one general practitioner, one clinical pharmacologist, one clinical pharmacist, one registered nurse, one home care nurse, and two community pharmacists. The expert panel were selected to reflect different settings and professional roles, and all 10

	members had at least 5 years of professional experience, held senior positions, and undertook regular clinical duties.
	Testing/validation: N/A
Saedder *et al.* (2016)[46]	**Summary of study:** Development of 'medicines risk score' (MERIS) to identify patients at increased risk of MEs.
	Study method: Literature search to identify variables, followed by two-round Delphi process to categorise medicines, and drug interactions as low, medium or high-risk[220]. Various scoring algorithms constructed, using theoretical weighting (which involved assigning importance to the various elements in the algorithm). Statistical testing used to compare the accuracy, sensitivity and specificity of various versions of the algorithm in three historic datasets, with adjustments made following each assessment. Final score validated in a prospective study of 53 patients. 'Detection limit' for each algorithm (i.e. the score to identify high-risk patients) was selected retrospectively to give the highest precision. Categorisation used for all variables in the risk score, but it is unclear how these categories were selected. Datasets ranged in size from 50 to 146 patients, providing 9-33 outcome events per group. The groups also differed in age, with two groups restricted to adults over the age of 65 years, whereas two included adult of all ages. There were also differences in: the number of medicines prescribed (those prescribed ≥ 8 medicines ranged from 40% to 92%); average number of MEs (from 0.2 to 1.1 per patient), and number of patients who experienced an ME (from 18% to 62%).
	Expert group: Panel of 36 experts (physicians and pharmacists) appointed by Danish Medical Societies, Danish Health and Medical Agency and the Danish Society of Pharmacists (no further details given). Survey administered electronically[220].
	Testing/validation: Final algorithm sensitivity 64%, specificity 75%, area under ROC curve 0.76 (95% CI 0.62-0.89).
Hickson *et al.* (2016)[51]	**Summary of study:** Assessment of a 'pharmaceutical assessment screening tool' (PAST), developed to assign a 'patient acuity level' (PAL) to patients (level 1, 2 or 3).
	Study method: Tool developed by consultant pharmacist in medication safety, based on similar tools in the literature[48 221 222] (including Cottrell's), but also includes patient-level[223 224] and medication-based risk factors based on high-risk medications known to cause serious harm[225]. A team of pharmacists piloted the tool to confirm face validity. Agreement on final tool and PAL sought using consensus methodology (no further details provided).
	Expert group: Junior and senior pharmacists from medical and surgical specialities within study site (900-bed teaching hospital in England). No further details given.
	Testing/validation: Quasi-experimental service evaluation to quantify agreement among pharmacist-documented and per-guidance PAL, no other testing or evaluation reported.

MRP = medication related problem, ME = medication error, ROC = receiver-operating characteristic, CI = confidence interval, N/A = not applicable

Appendix A5.1: Expert survey

Welcome to My Survey

Thank you for participating in this survey. Your feedback is important as we want to ensure that our research ideas are supported by experts, potential users, patients and the public. The survey should only take 4-5 minutes to complete. Please note that all responses to this survey will remain anonymous. We will not be tracking any identifying information.

Medicines optimisation is a key role for pharmacists, and our aim is to develop a prioritisation tool, the Medicines Optimisation Assessment Tool (MOAT), that can be used to target patients who are most in need of pharmacists' input during admission to hospital.

Predicting clinical risk is well established in healthcare, with prediction tools such as cardiac-risk calculators being used daily across the NHS. The Medicines Optimisation Assessment Tool will aim to predict which patients are at risk of experiencing a significant medication related problem (MRP) – defined as 'all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome (from prescribing to administration)'.

To predict which patients are at highest risk of medication related problems, we need to establish the relationship between risk factors and medication related problems (for example: "does risk increase with factors such as age and/or the number of medicines used?"). We have identified potential risk factors from published literature, but we need to make sure that the final list is comprehensive and clinically relevant – so would like the opinions of key groups of people (including pharmacy staff, doctors, nurses, researchers, patients and the public).

We would also appreciate your thoughts on how accurate the Medicines Optimisation Assessment Tool needs to be (known as the 'target sensitivity'). We appreciate that you may not feel able to comment on this, but any thoughts you have would be appreciated.

This research is being carried out as part of a National Institute for Health Research (NIHR) Clinical Doctoral Research Fellowship, and is a collaboration between University College London, the Luton and Dunstable University Hospital NHS Trust, and Watford General Hospital..

For more information please feel free to contact me.

Many thanks
Cathy

Cathy Geeson
NIHR Clinical Doctoral Research Fellow
Pharmacy Department
Luton and Dunstable University Hospital
Email: cathy.geeson@ldh.nhs.uk

Continued from previous page…

Review of potential risk factors

Listed below are the risk factors identified from published literature. Please review each risk factor and grade how important you believe it is in terms of causing medication related problems (MRPs).

As a reminder, medication related problems are defined as 'all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome (from prescribing to administration)'. This means that in addition to problems that can harm a patient (for example a dose which is too high or an allergic reaction) this also includes problems caused when treatment is 'sub-optimal' (for example if a patient is given a medicine which is unlikely to be effective, or a dose which is too low).

Continued from previous page…

1. **Patient related risk factors** Please rate each of the following possible risk factors in terms of importance (from very important to not important). If you do not feel able to comment then please tick the 'not sure' box.

	Very important	Important	50-50	Less important	Not important	**NOT SURE**
Age	○	○	○	○	○	○
Gender	○	○	○	○	○	○
Ethnicity	○	○	○	○	○	○
Swallowing difficulties	○	○	○	○	○	○
Allergies	○	○	○	○	○	○
Dependent living situation (e.g.nursing home / domiciliary care)	○	○	○	○	○	○
Whether the admission to hospital was elective (i.e. planned) or an emergency	○	○	○	○	○	○
Readmission to hospital within 30 days	○	○	○	○	○	○
Number of admissions to hospital in previous 6 months	○	○	○	○	○	○
Number of hospital outpatient visits in previous 6 months	○	○	○	○	○	○
Type of hospital department / medical speciality	○	○	○	○	○	○
Diagnosis / reason for admission	○	○	○	○	○	○
Number & details of co-morbidities (i.e. existing health problems)	○	○	○	○	○	○
Social deprivation	○	○	○	○	○	○

Continued from previous page…

2. **Medicine related risk factors**. Please rate each of the following possible risk factors in terms of importance (from very important to not important). If you do not feel able to comment then please tick the 'not sure' box.

Please note that we intend to collect information on medicines with are taken by mouth, injection, inhalation and rectally/vaginally and exclude medicines which are applied to the skin and eye drops.

	Very important	Important	50-50	Less important	Not important	**NOT SURE**
Number of medicines prescribed	○	○	○	○	○	○
Number of 'potentially inappropriate medicines' (these are medicines that are thought to have an increased risk of causing adverse effects in older people). Some healthcare professionals may know these as the 'STOPP tool' or 'Beers criteria'	○	○	○	○	○	○
Names of all medicines prescribed (to permit analysis by individual medicine, or by the type of medicine)	○	○	○	○	○	○
Route of medicines administration (e.g. by mouth, injection, inhalation, rectally or vaginally)	○	○	○	○	○	○
Dosing frequency for medication (i.e. how often the medicine is given)	○	○	○	○	○	○

STOPP = Screening Tool of Older People's potentially inappropriate Prescriptions

Continued from previous page…

3. **'Laboratory results' related risk factors** Please rate each of the following possible risk factors in terms of importance (from very important to not important). If you do not feel able to comment then please tick the 'not sure' box.

	Very important	Important	50-50	Less important	Not important	**NOT SURE**
Renal function (how well the kidneys work)	○	○	○	○	○	○
Liver function (how well the liver works)	○	○	○	○	○	○
Hyperlipidaemia (the amount of 'fat' and cholesterol in the blood)	○	○	○	○	○	○
White blood cell count (the cells in the blood that help fight infections)	○	○	○	○	○	○
Serum albumin level (a type of protein in the blood)	○	○	○	○	○	○
Serum sodium level (a routine blood test to monitor certain conditions such as dehydration)	○	○	○	○	○	○
Serum potassium level (a routine blood test to monitor certain conditions such as kidney disease)	○	○	○	○	○	○
Platelet count (part of the blood that helps the blood to clot)	○	○	○	○	○	○

4. We intend to collect information on the 'medicine related risk factors' for medicines which are taken by mouth, injection, inhalation, rectally or vaginally. This means we would not collect information on topical medicines (i.e. those applied to the skin and eye drops).

Do you believe that it is acceptable to exclude topical medicines?

○ YES

○ NO

○ UNSURE

Continued from previous page…

Are there any other possible risk factors?

We want to make sure that the final list of risk factors is comprehensive and clinically relevant - so please let us know whether there are any other factors that you think we should include.

5. Please list any other risk factors you believe should be considered for inclusion in this prediction tool.

6. Please give reasons for your suggestions given in question 5.

Continued from previous page…

How accurate should the MOAT be?

It is very rare for a prediction tool to be accurate 100% of the time, as there is a balance between being able to predict an outcome correctly, and overestimating how many people are at risk. The aim of this research is to produce a prediction tool which is able to accurately predict patients who are at risk of significant medication related problems 90% of the time (in other words be able to identify 9 out if 10 patients who are at risk). This is also known as 'target sensitivity'.

We chose a target of 90% as has been used by researchers carrying out similar work into medication safety (in terms of providing an acceptable level of safety for patients, while also being feasible to use in terms of the workload for pharmacists).

7. Do you believe that 90% is an acceptable target for the accuracy of the Medicines Optimisation Assessment Tool?

- ◯ YES
- ◯ NO
- ◯ UNSURE

8. If you would like to add any comments to explain your answer to question 4 please add them below.

Continued from previous page…

Your current role

9. Which of the following best describes your current role (you can choose more than one option)?

- [ ] Pharmacist / member of the pharmacy team
- [ ] Doctor
- [ ] Nurse
- [ ] Other healthcare professional
- [ ] Academic
- [ ] Patient or public representative
- [ ] Other (please specify)

Continued from previous page…

Would you like to receive updates?

10. If you would like to receive updates on our progress with this research please provide your contact details below:

Name	
Company	
Email Address	

Continued from previous page…

Thank-you!

Many thanks for completing this survey – your help is very much appreciated.

Please don't hesitate to get in touch if you would like any further information about the research project. My email address is cathy.geeson@ldh.nhs.uk

Best wishes
Cathy

Appendix A6.1: Comparison of comorbidity scales

System	Modified Cumulative Illness Rating Scale (CIRS)[128]	Charlson index[129]	Elixhauser system[130]
Cardiac	**Cardiac (heart only)** Includes: coronary arteries dx, heart failure, valvular heart dx, endocardities, miocardities, pericarditis, arrhythmias (extrasystoles, bundle branch blocks, AF, PMK placement), past MI, angina, acute coronary syndrome	**Myocardial infarct** *Excludes AF, valvular disease, arrhythmias & angina*	*Excluded: old myocardial infarction*
Heart failure	Included in cardiac	**Congestive heart failure**	**Congestive heart failure** Includes: rheumatic heart disease, various causes of heart failure, cardiomyopathy, left ventricular failure
Arrhythmias	Included in cardiac	*Excluded*	**Cardiac arrhythmias** Includes: heart block, arrhythmias (tachycardia, atrial fibrillation, bradycardia, ventricular fibrillation, cardiac pacemaker
Valvular disease	Included in cardiac	*Excluded*	**Valvular disease** Includes: cardiovascular syphilis, rheumatic valve diseases
Pulmonary vascular disorders	Included in vascular-haematopoietic	*Not included*	**Pulmonary vascular disorders** Includes: pulmonary embolism, pulmonary hypertension, pulmonary heart disease
Warfarin	*Not included*	**Use of warfarin** (additional component to predict cost)	*Not included*

Continued from previous page…

System	Modified Cumulative Illness Rating Scale (CIRS)	Charlson index	Elixhauser system
Vascular	**Vascular-haematopoietic** Artery disease: carotid atherosclerosis, peripheral arteries disease, aneurysms Venous disease: venous insufficiency, varices , deep vein thrombosis, pulmonary embolism, pulmonary hypertension Haematopoietic: anaemia, leucopenia, thrombocytopenia, haematological malignancy Lymphopoietic disease: chronic lymphatic oedema, lymphoma, spleen & thymus disease Immunologic disease: systemic lupus erythematosus, systemic sclerosis, sarcoidosis, hypersensitivity	**Peripheral vascular disease**	**Peripheral vascular disorders** Includes: atherosclerosis, aneurysms, peripheral vascular disease, cardiac & vascular implants
Hyperten -sion	**Hypertension**	**Hypertension** (additional component to predict cost)	**Hypertension**
Cerebro-vascular	Included in neurological	**Cerebrovascular disease** Includes: cerebrovascular accident, transient ischaemic attack	*Not included*
Dementia	Included in psychiatric / behavioural	**Dementia**	Included in specific comorbidities
Respiratory	**Respiratory** Includes: COPD, asthma, emphysema, restrictive pulmonary interstitial lung disease, malignancies of lung & pleura, pneumonia, smoking status	**Chronic pulmonary disease**	**Chronic pulmonary disease** Includes: pulmonary heart disease, bronchitis, emphysema, chronic obstructive pulmonary disease, asthma, bronchiectasis, pneumoconiosis
Ear, nose & throat	**Ear, nose, throat, larynx** Includes: glaucoma, cataracts, macular degeneration, otitis, dizziness, hearing impairment, rhinitis, pharyngitis, nasal polyps, sinusitis, dysphonia, laryngitis, malignancies	*Not included*	*Not included*

Continued from previous page…

System	Modified Cumulative Illness Rating Scale (CIRS)	Charlson index	Elixhauser system
Upper GI	**Upper gastrointestinal (GI)** Includes: intestinal tract from oesophagus to duodenum, & pancreatic tree (dysphagia, gastroesophageal reflux disease, hiatus hernia, oesophageal diverticula, gastritis, gastric / duodenal ulcer, pancreatitis, malignancies) Excludes diabetes	**Ulcer disease** *Excludes GI bleeding*	**Peptic ulcer disease** Includes: gastric, duodenal & gastrojejunal ulcer
Lower GI	**Lower GI** Includes: from small bowel to anus (Whipple's disease, diverticulosis, irritable bowel, malignancies, constipation)	*Excludes inflammatory bowel disease*	*Excludes inflammatory bowel disease*
Liver	**Hepatic** Includes: liver, gallbladder, biliary trees, portal system, hepatitis, cirrhosis, portal hypertension, malignancies	**Liver disease**	**Liver disease** Includes: viral & alcoholic hepatitis, varices, fatty liver, cirrhosis, hepatic failure, chronic liver disease
Endocrine / diabetes	**Endocrine-metabolic** Includes: diabetes, obesity, dyslipidaemia, thyroid disease; parathyroid disease, adrenal pathologies, hypogonadism, hypopituitarism, breast, malignancies of these glands. Also, electrolyte disorders, sepsis, systemic infections (tuberculosis, syphilis, AIDS), and poisonings (chronic by metals or acute by pesticides or carbon monoxide)	**Diabetes / diabetes with end organ damage** *Excludes hypopituitarism, adrenal insufficiency, recurrent acidosis*	**Diabetes** (uncomplicated & complicated scored separately) Includes: insulin dependent, non-insulin dependent, ketoacidosis, renal, ophthalmic, neurological, circulatory complications
Thyroid	Included in endocrine-metabolic	*Not included*	**Hypothyroidism**
Paralysis	Included under specific systems	**Hemiplegia**	**Paralysis** Includes: paraplegia, hemiplegia

Continued from previous page…

System	Modified Cumulative Illness Rating Scale (CIRS)	Charlson index	Elixhauser system
Renal	**Renal** Includes: kidney stones, renal failure, glomerulonephritis, nephrotic syndrome, pyelonephritis, nephropathy, carcinoma	**Moderate / severe renal disease**	**Renal failure** Includes: renal failure, chronic kidney disease' dialysis, renal transplant
Tumour	Included under specific systems	**Any tumour**	**Solid tumour without metastasis**
Leukae-mia	Included in vascular-haematopoietic	**Leukaemia**	Included under lymphoma
Lymph-oma	Included in vascular-haematopoietic	**Lymphoma**	**Lymphoma**
Metastat-ic cancer	Included under specific systems	**Metastatic solid tumour**	**Metastatic cancer**
AIDS / HIV	Included in endocrine-metabolic	**AIDS / HIV**	**AIDS / HIV**
Neurological	**Neurological** Includes: stroke, neurodegenerative diseases (Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis etc), myelopathies, trauma with neurological outcomes, epilepsy, neuropathies, primary tumours, migraines, insomnia *Excludes dementia*	*Excludes Parkinson's disease, seizures, syncope*	**Other neurological disorders** Includes: Huntington disease, cerebellar ataxia, motor neurone disease, Parkinson's disease, multiple sclerosis, epilepsy, encephalopathy *Excludes: cerebrovascular disease*

Continued from previous page…

System	Modified Cumulative Illness Rating Scale (CIRS)	Charlson index	Elixhauser system
Musculoskeletal	**Musculoskeletal-intergumentary** Includes: osteoarthritis, osteoporosis, bone fracture, neoplasm, rheumatoid arthritis, polymyalgia rheumatic, muscular injuries, pressure sores, dermatological disease	**Connective tissue disease**	**Rheumatoid arthritis / collagen vascular diseases** Includes: scleroderma, rheumatoid arthritis, ankylosing spondylitis, systemic lupus erythematosus, polymyositis, Sjögren syndrome, polymyalgia rheumatic, ankylosing spondylitis *Excludes: osteoarthritis*
Skin	Included in musculoskeletal-intergumentary	**Skin ulcers / cellulitis** (additional component to predict cost)	*Excluded*
Coagulopathy	Included in vascular-haematopoietic	*Excluded*	**Coagulopathy** Includes: defibrination syndrome, factor VIII & IX, XI deficiency, thrombophilia, Idiopathic thrombocytopenic purpura, thrombocytopenia
Obesity	Included in endocrine-metabolic	*Not included*	**Obesity**
Weight loss	Not included	*Not included*	**Weight loss**
Fluid & electrolytes	Included in endocrine-metabolic	*Not included*	**Fluid & electrolyte disorders** Includes: dehydration, hypovolaemia, hypo-osmolality and hyponatraemia, acidosis, alkalosis, hyperkalaemia, hypokalaemia, fluid overload

Continued from previous page…

System	Modified Cumulative Illness Rating Scale (CIRS)	Charlson index	Elixhauser system
Anaemia	Included in vascular-haematopoietic	*Not included*	**Blood loss anaemia**
	Included in vascular-haematopoietic	*Not included*	**Deficiency anaemia** Includes: iron deficiency anaemia, vitamin B_{12} deficiency anaemia, folate deficiency anaemia, other megaloblastic anaemia
Alcohol abuse	*Not included*	*Not included*	**Alcohol abuse** Includes: harmful alcohol use, and complications of alcohol abuse (e.g. polyneuropathy)
Drug abuse	*Not included*	*Not included*	**Drug abuse**
Psychiatric	**Psychiatric / behavioural** Includes dementia, depression, anxiety, agitation / delirium, psychosis	*Not included*	**Psychoses** Includes: schizophrenia, delusional disorder, schizoaffective disorder, bipolar affective disorder
Depress-ion	Included in psychiatric / behavioural	**Depression** (additional component to predict cost)	**Depression**
Genito-urinary	**Other genitourinary** Includes: ureters, bladder, urethra, prostate, genitals, uterus, ovaries	*Not included*	*Not included*

AIDS = acquired immunodeficiency syndrome, HIV = human immunodeficiency virus

Appendix A6.2: Grouping used for MOAT comorbidity count

Grouping of comorbidities (based on the categories used in the Modified Cumulative Illness Rating Scale)[128]
Cardiac 1. Ischaemic heart disease (including old myocardial infarction, angina pectoris, aortocoronary bypass graft, presence of coronary angioplasty implant and graft) 2. Arrhythmias (including atrial fibrillation, presence of electronic cardiac devices, ventricular fibrillation, heart block) 3. Heart failure 4. Valvular heart disease 5. Other (including. septal defects, cardiomyopathy, cardiomegaly, myocarditis)
Hypertension 6. Hypertension
Vascular-haematopoietic 7. Artery disease (including atherosclerosis, peripheral arteries disease) 8. Aneurysms 9. Venous disease (including venous insufficiency, oesophageal varices , deep vein thrombosis, pulmonary embolism, pulmonary hypertension) 10. Haematopoietic (including anaemia, leucopenia, thrombocytopenia) 11. Lymphopoietic disease (including lymphadenitis, lymphoedema, diseases of spleen) 12. Immunologic disease (including systemic lupus erythematosus, systemic sclerosis, sarcoidosis) 13. Raynaud syndrome
Respiratory 13. Chronic obstructive pulmonary disease 14. Asthma 15. Interstitial lung disease 16. Other (including sleep apnoea, absence of part of lung)
Eye, ear, nose & throat 17. Glaucoma 18. Retinal disease (including blindness, retinopathy, macular degeneration) 19. Disorders of balance (including Meniere's disease, vertigo) 20. Chronic sinusitis
Upper gastrointestinal (GI) 21. Oesophageal (including gastro-oesophageal reflux disease, oesophagitis) 22. Stomach / duodenum (including gastritis, duodenitis) 23. Pancreatic tree (including pancreatitis, diseases of pancreas *excluding diabetes*) 24. Dyskinesia of oesophagus 25. Other (including absence of part of stomach, tracheo-oesophageal fistula)
Lower GI 26. Inflammatory bowel disease 27. Coeliac disease / diverticulitis / irritable bowel syndrome 28. Other (including Meckel diverticulum, vesicointestinal fistula)
Hepatic 29. Liver (including hepatitis, hepatic failure, cirrhosis) 30. Gall bladder (including cholecystitis, obstruction of bile duct, biliary cirrhosis) 31. Portal hypertension

Continued from previous page…

Renal
32. Anatomical (including absence of kidney, cysts, calculus)
33. Functional (including chronic kidney disease, renal failure)
34. Other (including hydronephrosis, nephrotic syndrome, pyonephrosis)
Other genitourinary
35. Prostatic hypertrophy
36. Bladder (including calculus, cystocele)
37. Ovaries / uterus (including leiomyoma, polycystic ovary disease)
Musculoskeletal-intergumentary
38. Osteoarthritis (including arthrosis, spondylosis, spinal stenosis)
39. Rheumatoid arthritis / fibromyalgia / psoriasis / pemphigoid / myasthenia gravis
40. Gout / crystal arthropathies
41. Muscular dystrophy
42. Nummular dermatitis
43. Osteogenesis imperfect
44. Osteoporosis
45. Ehlers-Danlos syndrome
Neurological
46. Stroke
47. Parkinson's disease
48. Other degenerative diseases (including multiple sclerosis, motor neurone disease)
49. Neuropathies
50. Epilepsy
51. Myelopathy / dorsalgia / sciatica
52. Paralysis
53. Hydrocephalus
54. Huntington disease
55. Neuropathic arthropathy
Endocrine-metabolic
56. Diabetes
57. Thyroid disorders
58. Dyslipidaemia
59. Parathyroid disorders
60. Adrenal disease
61. Systemic infectious diseases (human immunodeficiency virus, syphilis)
62. Disorders of pituitary gland
63. Breast disorders
64. Syndrome of inappropriate secretion of antidiuretic hormone
65. Paget's disease
66. Obesity
Psychiatric / behavioural
67. Dementia
68. Depression / anxiety / bipolar affective disorder / panic disorder
69. Psychosis / schizophrenia
70. Agoraphobia / eating disorders
71. Autism
72. Developmental disorders
73. Postviral fatigue syndrome / somataform disorder
Tumours / malignancies
74. Benign neoplasms
75. Malignant neoplasm
76. Leukaemia / lymphoma

Appendix A6.3: Conventions for medicine data collection

Scenario	Hospital A (electronic prescribing records)	Hospital B (handwritten prescribing records)
Medicine prescribed regularly, but as intravenous **or** oral	Record as 'IV or oral'	As Hospital A
Medicine prescribed regularly, but patient 'refused'	Record on the spread sheet irrespective of whether doses actually administered (i.e. the medicine had been prescribed)	As Hospital A
Variable doses prescribed (e.g. codeine 30mg or 60mg)	Only list once	As Hospital A
Intravenous fluids	Record as 'IV FLUIDS' (i.e. not individually) i.e. binary variable (yes/no)	As Hospital A If a medicine is prescribed by infusion – this needs to be recorded separately, e.g. omeprazole & aminophylline are always given by infusion
Intravenous / subcutaneous fluids given once only	Treat as 'once only' dose therefore don't include	As Hospital A
Medicine prescribed then stopped & re-prescribed the same day (different dose)	Only list once	As Hospital A
Medicine prescribed but no doses given	Record (i.e. the medicine had been prescribed)	As Hospital A If only partly prescribed (i.e. 'crossed off' before fully prescribed) assume prescriber changed their mind before finishing the prescription - & do not record
Medicine 'suspended' (i.e. prescribed but withheld)	Record (i.e. the medicine had been prescribed)	As Hospital A
Drug prescribed to start on a date after the patient was discharged	Ignore	As Hospital A
Medicine prescribed twice (different frequency, route, formulation / administration device or dose)	Record both	As Hospital A
Start date on day after the first full day of admission, but at 8am	Include in medicine count as prescribed on first day of admission	As Hospital A
Bedtime snack	Ignore	As Hospital A
Medicines prescribed on day after first day of admission but 'once only' dose was prescribed on the first day	Include in medicine count as prescribed on first day of admission	As Hospital A

Appendix A6.4: Diagnoses included in MOAT categories for primary diagnosis

MOAT category	Diagnoses included
Cardiovascular system	Valvular disease; hypertension; angina; myocardial infarction; ischaemic heart disease; atherosclerotic heart disease; aneurysm; pericarditis; arrhythmias; cardiac arrest; heart failure; atherosclerosis of extremities; phlebitis; varicose veins; oesophageal varices; orthostatic hypotension; pulmonary embolism; primary pulmonary hypertension; tachycardia; bradycardia; haematopoietic disorders (e.g. anaemia, haematological malignancies); lymphopoietic disease (e.g. lymphoma, enlarged lymph nodes); immunologic disease (e.g. hypergammaglobulinaemia).
Respiratory system	Tracheitis; pneumonia; bronchitis; emphysema; chronic obstructive pulmonary disease; asthma; bronchiectasis; pneumonitis; interstitial pulmonary disease; pyothorax; pneumothorax; respiratory failure; pulmonary infections (e.g. tuberculosis, mycobacteria); neoplasms; mesothelioma of pleura; haemoptysis; dyspnoea.
Gastrointestinal system	Diseases of tongue; tonsillitis; oesophagitis; achalasia of cardia; gastro-oesophageal reflux disease; ulcers of gastrointestinal tract; gastritis / duodenitis; obstruction; appendicitis; hernias; inflammatory bowel disease; gastroenteritis; paralytic ileus; diverticular disease; constipation; haemorrhoids; peritonitis; liver disease; disease of gallbladder and bile ducts; pancreatic disease; hematemesis; enteritis / enterocolitis; intestinal viral infections; tuberculosis of intestines; stomatitis; neoplasms; dysphagia.
Genitourinary system	Nephritis; renal failure; pyonephrosis; calculus (kidney & urinary); cystitis; urinary tract infection; hyperplasia of prostate; diseases of testis, uterus, ovaries, vagina; haematuria; urine retention; neoplasms.
Musculoskeletal -intergumentary systems	Rheumatoid arthritis; arthropathies; gout; painful joints; spondylosis; myelopathies; cervicalgia; sciatica; neoplasm of bone; neoplasms of skin; abscess / carbuncle; cellulitis; lichen simplex chronicus; nail disorders; psoriasis; urticaria; skin ulcers.
Endocrine-metabolic diseases	Diabetes; thyroid disorders; parathyroid disease; syndrome of inappropriate secretion of antidiuretic hormone; pituitary disease; adrenal disease; electrolyte disorders; disorders of plasma-protein metabolism; systemic infections (sepsis, malaria); breast cancer; neoplasms.
Nervous system and mental disorders	Meningitis / encephalomyelitis; Parkinson's disease; Alzheimer's disease; degenerative diseases of nervous system; multiple sclerosis; epilepsy; migraine; cluster headaches; amnesia; transient ischaemic attack; Bell's palsy; neuropathies; encephalopathy; myelopathies; other dementias / delirium / cognitive disorders; brain / nervous system related pathologies (e.g. infections and neoplasms); stroke; paraesthesia; depression; anxiety; bipolar disorder; eating disorders somatoform disorders.
Other	Varicella; infectious mononucleosis; mesothelioma / neoplasms (where site not specified); amyloidosis; effects of alcohol use / abuse; eye and vestibular disorders; poisoning; allergy / anaphylaxis; non-specific symptoms (e.g. abnormal investigative findings, cough, precordial / chest pain, abdominal pain, nausea and vomiting, dizziness, fever, headache, malaise and fatigue, somnolence, disorientation, epistaxis); non-specific findings (e.g. haemorrhage - site not specified, difficulty walking, tendency to fall, syncope, localised swelling, collapse, volume depletion); superficial injuries (fractures, traumatic injuries).

Appendix A6.5: Medication related problem data collection form

MOAT Study – MRP Data Collection Form

Please complete 1 form per MRP identified unless there are a number of MRPs of the **same category** for the **same patient** (e.g. more than 1 omission found at MR, dosage adjustment needed due to renal impairment). For these please number & list each MRP on the same form with details of the medicines involved.

Patient details							
Study code:		**Hospital No:**		**Ward:**			
MRP details							
Date MRP occurred:	If MRP occurred before admission please write 'pre-admission'	**When MRP identified:**	During 'ward' duties	During 'departmental' duties	Other (e.g. referrals, ward round)		
Date MRP resolved:		**Who resolved MRP:**	Pharmacy staff		Other HCP		
Name of drug(s) involved & route: (if applicable):		**Stage in admission MRP identified:**	**During first ward review** (or before)	**Remainder of inpatient stay** (after initial full validation)	**TTA screening**		
					MRP that occurred when TTA written	MRP occurred prior to TTA writing	
		Is it a Meds Rec discrepancy?	Yes		No		
Brief details of MRP (please record sufficient details to permit the potential severity of the MRP to be assessed):							
Do you consider that this MRP could have been prevented? **Select 'Yes' if 'preventable'** i.e. a medication error **Select 'No' if 'non-preventable'** i.e. it could **not** have been predicted e.g. an ADR						Yes	No
Name of person completing form ***(please also indicate role):***		**Pharmacist**	**MMPT**	**Other** (e.g. disp staff/ pre-reg, ward band3/4)			

Classification of MRP **Please tick 1 box only – if unable to categorise please discuss with Cathy (Telephone: 07949 725540)**

Tips:

- MRPs are: 'all circumstances involving a patient's drug treatment that actually, or potentially, interferes with the achievement of an optimal outcome'
- If the cause of the MRP is about **drug prescription**, see categories 1-4 first
- If the cause is about the way the **drug is used or administered**, see category 5
- For **supply, incomplete prescription or selection errors** see category 6
- For **monitoring** issues see category 7
- If an **adverse reaction** has occurred, or **no obvious reason for treatment failure**, see category 8

Medicines Reconciliation discrepancies – please classify as follows:

- drug not prescribed (unintentionally) – select 1.4 (missing therapy)
- drug prescribed when no longer taken by patient (unintentionally) – select 1.2 (no indication for drug)
- incorrect dose or frequency – see category 3 (dose selection)

MRP = medication related problem, Meds Rec = medicines reconciliation,
TTA = discharge prescription ("To Take Away"),
MMPT = medicines management pharmacy technician

Continued from previous page…

1. Drug selection		3. Dose selection	
1.1 Inappropriate drug e.g. precaution/contraindication/allergy; ineffectiveness (either drug not effective, or not effective for the indication being treated; safer/more appropriate alternative available; condition refractory to drug)		**3.1 Drug dose too low** i.e. overall daily dose	
		3.2 Drug dose too high i.e. overall daily dose	
		3.3 Dosage regimen not frequent enough i.e. total daily dose OK but dosing less frequent than recommended (e.g. trimethoprim for UTI prescribed as 400mg daily when standard dose is 200mg BD)	
1.2 No indication for drug / duplication e.g. inappropriate duplication of therapy; too many drugs for an indication; no documented indication; drug not currently indicated/ not currently taken; or indication does not warrant drug treatment			
		3.4 Dosage regimen too frequent i.e. total daily dose OK but dosing more frequent than recommended (e.g. trimethoprim for UTI prescribed as 100mg QDS when standard dose is 200mg BD)	
1.3 Interaction (drug-drug, or drugs and food/alcohol) i.e. any type of drug interaction			
1.4 Indication not treated/missing therapy e.g. anticoagulant needed due to stroke risk; history of osteoporosis but no treatment prescribed; Meds Rec discrepancy – drug omitted		**3.5 Dose needs adjustment to organ function or change in disease state** i.e. drug still suitable & the dose *initially* prescribed was OK (e.g. renal function alteration requiring dose modification; pain resolving requiring analgesic dose to be reviewed)	
1.5 More cost effective drug available i.e. MRP not related to safety or efficacy – just cost			
1.6 Synergistic/preventive drug required and not given e.g. probiotics not prescribed for high-risk patient on broad spectrum antibiotics; additional antihypertensive needed to control blood pressure		**4. Treatment duration / withdrawal**	
		4.1 Duration of treatment too short e.g. stat dose of an antibiotic prescribed when 3 day course needed	
2. Drug form		**4.2 Duration of treatment too long** e.g. antibiotic course too long	
2.1 Inappropriate or suboptimal drug form e.g. tablets prescribed for a patient unable to swallow; inhaler prescribed for patient with poor dexterity; IV treatment prescribed when oral more appropriate (e.g. quinolone)		**4.3 Abrupt withdrawal** i.e. withdrawal syndrome due to abrupt discontinuation of therapy	
5. Drug use process		**6. Logistics**	
5.1 Inappropriate timing of administration /dosing by prescriber; administration error by nurse e.g. diuretics prescribed at night; nitrates prescribed without a nitrate free period; dose rounding required; IV injection/infusion administered too quickly; incorrect dose administered		**6.1 Prescribed drug not available** e.g. non-formulary/out-of-stock	
5.2 Drug underused/under-administered e.g. patient refuses some (not all) doses; some (not all) doses not administered by nursing staff (despite drug being on ward)		**6.2 Drug order incorrect, incomplete, poorly legible/illegible/illegal/incorrect/allergy status incomplete (i.e. not all allergies recorded)** e.g. dose/strength/route/form etc missing from prescription; poor legibility; CD prescription not legal; incorrect (or no) route selected (e.g. IV bolus vs injection)	
5.3 Drug overused/over-administered e.g. patient overuses a drug at the bedside such as an inhaler/cream; nurse administers more frequent doses than prescribed		**6.3 Error in drug selection** e.g. doctor selects wrong drug (on prescribing software or from memory); pharmacy dispense wrong drug; nurse administers wrong drug, or wrong patient's drugs; patient takes another patient's drugs	
		7. Monitoring	
5.4 Drug not taken/administered at all e.g. patient refuses to take the drug; nurse does not administer drug (all doses) despite it being on ward		**7.1 Monitoring too frequent** e.g. TDM done too soon after dose change, once-daily gentamicin monitoring done daily	
5.5 Wrong drug taken by patient e.g. patient selects the wrong inhaler at bedside		**7.2 No or too infrequent monitoring** e.g. TDM not done/done less than recommended; patient on lithium & thyroid function not checked	
5.6 Drug abused i.e. suspected overuse use of a drug by a patient e.g. request for repeated opiate/cyclizine administration		**7.3 Inappropriate test ordered** e.g. INR checked for a patient on NOAC	
5.7 Patient or nurse uses drug incorrectly through lack of knowledge or barriers (e.g swallowing, dexterity) e.g. patient chews/nurse crushes SR tablet; patient unable to use inhaler device due to dexterity problems (prescriber unaware); PV drug accidentally used PR; eye drops accidentally used nasally		**8. Unexpected reaction /ADR / no obvious cause**	
		8.1 An ADR occurred i.e. a reaction which could **not** have been anticipated (i.e. not dose related; occurred at normal therapeutic dose; no precautions/contraindications present) e.g. allergic drug reaction in patient with NKDA; injection site reaction	
5.9 Drugs stored inappropriately/ expired drug administered/preparation error e.g. incorrect storage of heat/light/moisture labile drugs; expired drug administered; drug incorrectly reconstituted or incorrect diluent/IV fluid used		**8.2 No obvious cause of treatment failure** i.e. no response despite optimal treatment e.g. optimal dose/drug prescribed & condition not thought to be refractory to treatment	

IV = intravenous, SR = sustained release, CD = controlled drug,
TDM = therapeutic drug monitoring, INR = International Normalised Ratio,
NOAC = novel oral anticoagulant, PV = per vagina, PR = per rectum,
ADR = adverse drug reaction, UTI = urinary tract infection, BD = *bis in die* (twice daily),
QDS = *quarter die sumendum*' (four times daily)

Appendix A6.6: Breakdown of missing data by study variable

Variable	Admissions with missing data		
	Hospital A (admissions = 1,006) n (% of admissions)	**Hospital B (admissions = 497)** n (% of admissions)	**All patients (admissions = 1,503)** n (% of admissions)
Medicines reconciliation completed*	0	0	0
Age	0	0	0
Gender*	0	0	0
Socioeconomic status	3 (0.3)	3 (0.6)	6 (0.4)
Ethnic origin*	62 (6.2)	34 (6.8)	96 (6.4)
Previous allergy	1 (0.1)	0	1 (0.07)
Body mass index	257 (25.6)	84 (16.9)	341 (22.7)
Number of hospital admissions in previous 6 months	0	0	0
Primary diagnosis	0	0	0
Number of comorbidities	0	0	0
History of dementia	0	0	0
Length of hospital stay*	0	0	0
Number of medicines	0	0	0
High-risk medicines use	0	0	0
Parenteral administration route	0	0	0
Renal function	6 (0.6)	3 (0.6)	9 (0.6)
Liver disease	0	0	0
Serum albumin	7 (0.7)	19 (3.8)	26 (1.7)
Serum potassium	20 (2.0)	10 (2.0)	30 (2.0)
Serum sodium	3 (0.3)	0	3 (0.2)
White cell count	5 (0.5)	1 (0.2)	6 (0.4)
Platelet count	6 (0.6)	2 (0.4)	8 (0.5)

* Data collected for descriptive purposes and not included in the regression analyses to develop the MOAT

Appendix A6.7: Characteristics of admissions with missing values and completely observed data

Characteristic	Admissions with missing data = 449 n (% of admissions with missing data)	Admissions with complete data = 1,054 n (% of admissions with complete data)	*p* value (test for difference between admissions with and without missing data)
Medication related problem (MRP) occurrence*			
Number of admissions with outcome event (at least one moderate or severe preventable MRP)	140 (31.2)	470 (44.6)	<0.0001 (Chi-square)
Patient related			
Age (years)	Median: 71 IQR: 56-83	Median: 77 IQR: 60-86	<0.001 (Mann-Whitney)
Gender (female)	179 (39.9)	514 (48.8)	0.0015 (Chi-square)
Socioeconomic status†‡	Median: 50.6 IQR: 27.6-78.0	Median: 49.5 IQR: 31.2-80.0	0.384 (Mann-Whitney)
Ethnic origin (White)†	292 (83.0)	915 (86.8)	0.075 (Mann-Whitney)
Body mass index† (kg/m^2)	Median: 26.8 IQR: 22.5-29.5	Median: 24.7 IQR: 21.4-29.0	0.029 (Mann-Whitney)
Number of hospital admissions in previous 6 months	Median: 0 IQR: 0-1	Median: 0 IQR: 0-1	<0.001 (Mann-Whitney)
Number of comorbidities	Median: 3 IQR: 1-4	Median: 4 IQR: 2-6	<0.001 (Mann-Whitney)
History of dementia	54 (12.0)	107 (10.2)	0.282 (Chi-square)
Length of hospital stay (days)	Median: 3 IQR: 1-6	Median: 7 IQR: 3-15	<0.001 (Mann-Whitney)
Medicines related			
Number of medicines	Median: 6 IQR: 4-10	Median: 8 IQR: 5-11	<0.001 (Mann-Whitney)
Parenteral administration	253 (56.3)	755 (71.6)	<0.001 (Mann-Whitney)
Laboratory results			
Renal function† - estimated glomerular filtration rate§ ≥ 60 $ml/min/1.73m^2$	313 (71.1)	690 (65.5)	0.0334 (Chi-square)
Serum albumin within standard reference †	201 (47.5)	420 (39.9)	0.0069 (Chi-square)
Serum potassium within standard reference range†	372 (88.8)	928 (88.3)	0.692 (Chi-square)
Serum sodium within standard reference range†	377 (84.5)	868 (82.4)	0.305 (Chi-square)
White cell count within standard reference range†	274 (61.9)	618 (58.6)	0.237 (Chi-square)
Platelet count within standard reference range†	385 (87.3)	900 (85.4)	0.332 (Chi-square)

* MRPs considered by expert panel to be true MRPs

† For patients without missing data (further details provided in Appendix A6.6)

‡ Ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]. Deprivation rank based on patients' postcode, shown as the ranked position as a percentage of all neighbourhoods in England (where 1 is the most deprived)

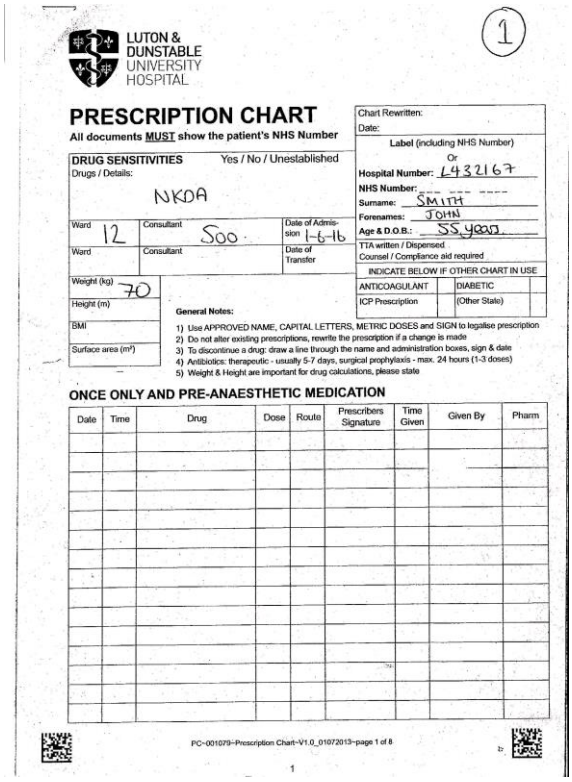
§ Glomerular filtration rate estimated using modified Modification of Diet in Renal Disease (MDRD) equation[121]. Glomerular filtration rate cut-off of < 60 $ml/min/1.73m^2$ used as an indicator of chronic kidney disease[125]

IQR = interquartile range

Bonferroni adjusted *p* value used to judge statistical significance 0.0028 (based on 18 statistical tests)

Appendix A7.1: Medication charts used for pharmacist medication related problem identification exercise

Medication chart 1*:



LUTON & DUNSTABLE UNIVERSITY HOSPITAL

PRESCRIPTION CHART

All documents MUST show the patient's NHS Number

Chart Rewritten:
Date:

Label (including NHS Number)
Or
Hospital Number: L432167
NHS Number:
Surname: SMITH
Forenames: JOHN
Age & D.O.B.: 55 years.
TTA written / Dispensed
Counsel / Compliance aid required

DRUG SENSITIVITIES Yes / No / Unestablished
Drugs / Details: NKDA

Ward 12 Consultant SOO. Date of Admission 1-6-16
Ward Consultant Date of Transfer

Weight (kg) 70
Height (m)
BMI
Surface area (m²)

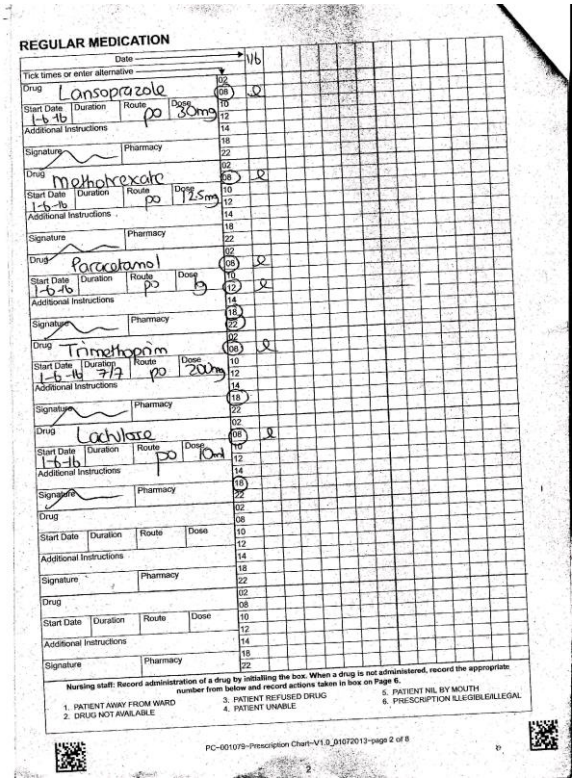
INDICATE BELOW IF OTHER CHART IN USE
ANTICOAGULANT DIABETIC
ICP Prescription (Other State)

General Notes:
1) Use APPROVED NAME, CAPITAL LETTERS, METRIC DOSES and SIGN to legalise prescription
2) Do not alter existing prescriptions, rewrite the prescription if a change is made
3) To discontinue a drug: draw a line through the name and administration boxes, sign & date
4) Antibiotics: therapeutic - usually 5-7 days, surgical prophylaxis - max. 24 hours (1-3 doses)
5) Weight & Height are important for drug calculations, please state

ONCE ONLY AND PRE-ANAESTHETIC MEDICATION

Date	Time	Drug	Dose	Route	Prescribers Signature	Time Given	Given By	Pharm

PC-001079–Prescription Chart–V1.0_01072013–page 1 of 8



REGULAR MEDICATION

Date 1/6
Tick times or enter alternative

Drug Lansoprazole — Start Date 1-6-16 — Duration — Route po — Dose 30mg — Additional Instructions — Signature — Pharmacy — 08 ✓

Drug Methotrexate — Start Date 1-6-16 — Duration — Route po — Dose 12.5mg — Additional Instructions — Signature — Pharmacy — 08 ✓

Drug Paracetamol — Start Date 1-6-16 — Duration — Route po — Dose 1g — Additional Instructions — Signature — Pharmacy — 08 ✓ 12 ✓

Drug Trimethoprim — Start Date 1-6-16 — Duration 7/7 — Route po — Dose 200mg — Additional Instructions — Signature — Pharmacy — 08 ✓

Drug Lactulose — Start Date 1-6-16 — Duration — Route po — Dose 10ml — Additional Instructions — Signature — Pharmacy — 08 ✓

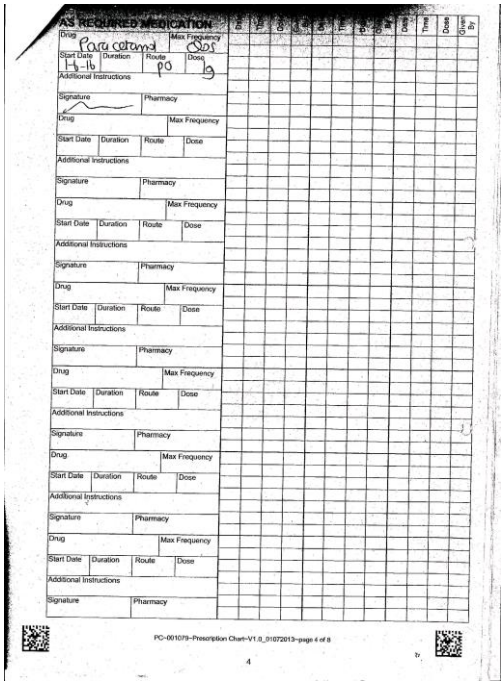
Drug — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Nursing staff: Record administration of a drug by initialling the box. When a drug is not administered, record the appropriate number from below and record actions taken in box on Page 6.
1. PATIENT AWAY FROM WARD
2. DRUG NOT AVAILABLE
3. PATIENT REFUSED DRUG
4. PATIENT UNABLE
5. PATIENT NIL BY MOUTH
6. PRESCRIPTION ILLEGIBLE/ILLEGAL

PC-001079–Prescription Chart–V1.0_01072013–page 2 of 8



AS REQUIRED MEDICATION

Drug Paracetamol — Max Frequency QDS — Start Date 1-6-16 — Duration — Route po — Dose 1g — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

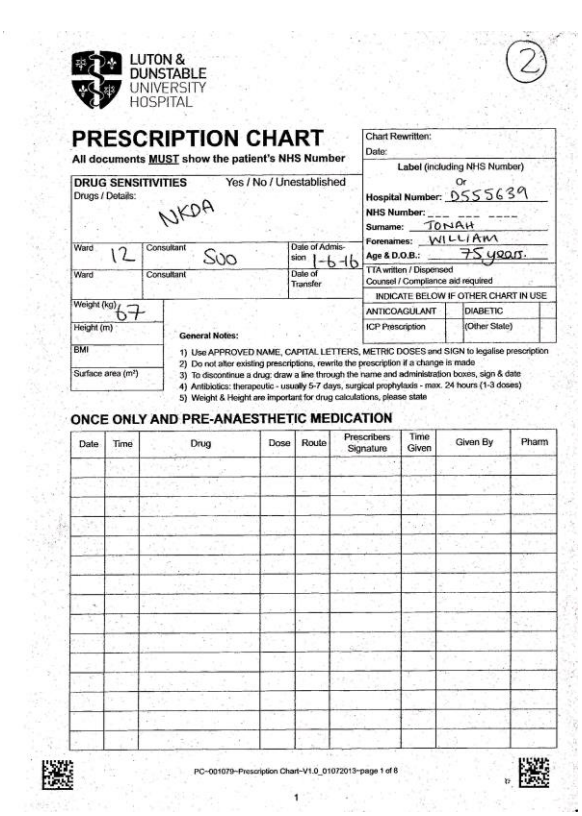
Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

Drug — Max Frequency — Start Date — Duration — Route — Dose — Additional Instructions — Signature — Pharmacy

PC-001079–Prescription Chart–V1.0_01072013–page 4 of 8

* Actual paperwork used was A4

Medication chart 2:



LUTON & DUNSTABLE UNIVERSITY HOSPITAL

2

PRESCRIPTION CHART

All documents MUST show the patient's NHS Number

Chart Rewritten:
Date:
Label (including NHS Number)
Or
Hospital Number: 0555639
NHS Number:
Surname: TONAH
Forenames: WILLIAM
Age & D.O.B.: 75 years.
TTA written / Dispensed
Counsel / Compliance aid required
INDICATE BELOW IF OTHER CHART IN USE
ANTICOAGULANT | DIABETIC
ICP Prescription | (Other State)

DRUG SENSITIVITIES Yes / No / Unestablished
Drugs / Details: NKDA

Ward: 12 | Consultant: SDO | Date of Admission: 1-6-16
Ward | Consultant | Date of Transfer

Weight (kg): 67
Height (m)
BMI
Surface area (m²)

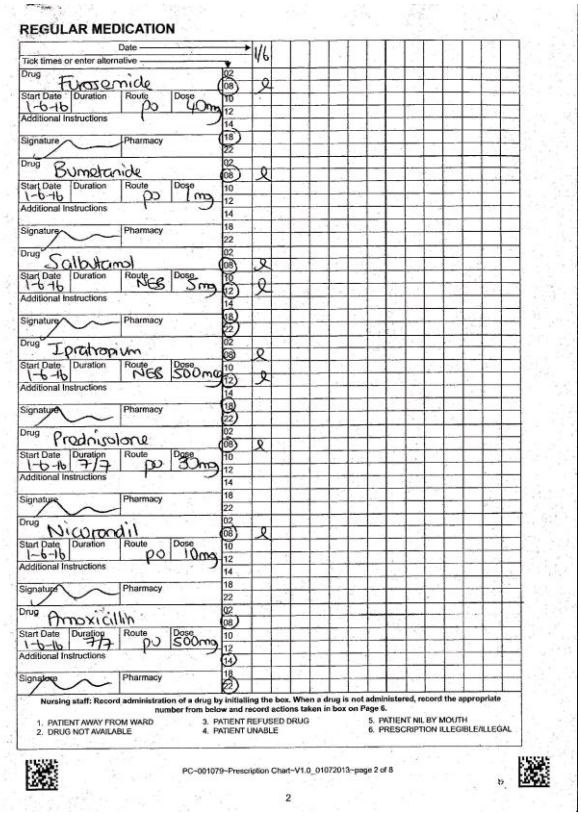
General Notes:
1) Use APPROVED NAME, CAPITAL LETTERS, METRIC DOSES and SIGN to legalise prescription
2) Do not alter existing prescriptions, rewrite the prescription if a change is made
3) To discontinue a drug: draw a line through the name and administration boxes, sign & date
4) Antibiotics: therapeutic - usually 5-7 days, surgical prophylaxis - max. 24 hours (1-3 doses)
5) Weight & Height are important for drug calculations, please state

ONCE ONLY AND PRE-ANAESTHETIC MEDICATION

Date	Time	Drug	Dose	Route	Prescribers Signature	Time Given	Given By	Pharm

PC-001079-Prescription Chart-V1.0_01072013-page 1 of 8

1



REGULAR MEDICATION

Date → 1/6
Tick times or enter alternative →

Drug: Furosemide
Start Date: 1-6-16 | Duration | Route: po | Dose: 40mg
Additional Instructions
Signature | Pharmacy

Drug: Bumetanide
Start Date: 1-6-16 | Duration | Route: po | Dose: 1mg
Additional Instructions
Signature | Pharmacy

Drug: Salbutamol
Start Date: 1-6-16 | Duration | Route: NEB | Dose: 5mg
Additional Instructions
Signature | Pharmacy

Drug: Ipratropium
Start Date | Duration | Route: NEB | Dose: 500mg
Additional Instructions
Signature | Pharmacy

Drug: Prednisolone
Start Date: 1-6-16 | Duration: 7/7 | Route: po | Dose: 30mg
Additional Instructions
Signature | Pharmacy

Drug: Nicorandil
Start Date: 1-6-16 | Duration | Route: po | Dose: 10mg
Additional Instructions
Signature | Pharmacy

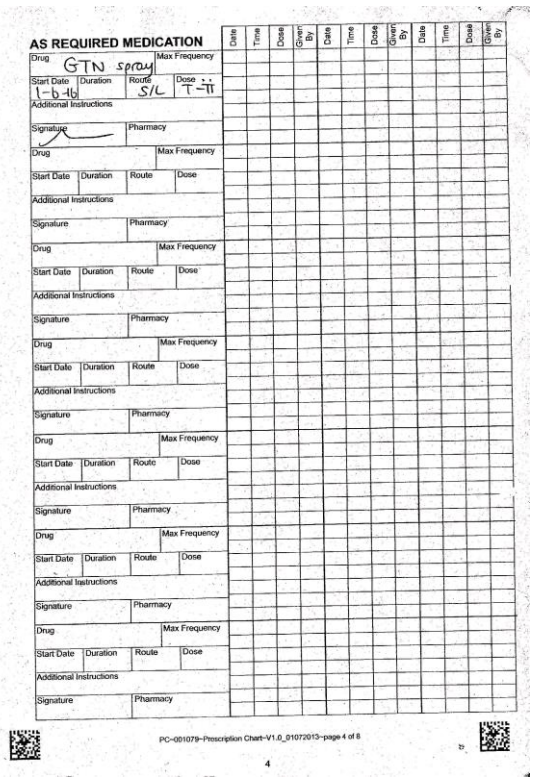
Drug: Amoxicillin
Start Date: 1-6-16 | Duration: 3/7 | Route: po | Dose: 500mg
Additional Instructions
Signature | Pharmacy

Nursing staff: Record administration of a drug by initialling the box. When a drug is not administered, record the appropriate number from below and record actions taken in box on Page 6.

1. PATIENT AWAY FROM WARD
2. DRUG NOT AVAILABLE
3. PATIENT REFUSED DRUG
4. PATIENT UNABLE
5. PATIENT NIL BY MOUTH
6. PRESCRIPTION ILLEGIBLE/ILLEGAL

PC-001079-Prescription Chart-V1.0_01072013-page 2 of 8

2



AS REQUIRED MEDICATION

Drug: GTN spray | Max Frequency
Start Date: 1-6-16 | Duration | Route: SL | Dose: 1-2
Additional Instructions
Signature | Pharmacy

Drug | Max Frequency
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

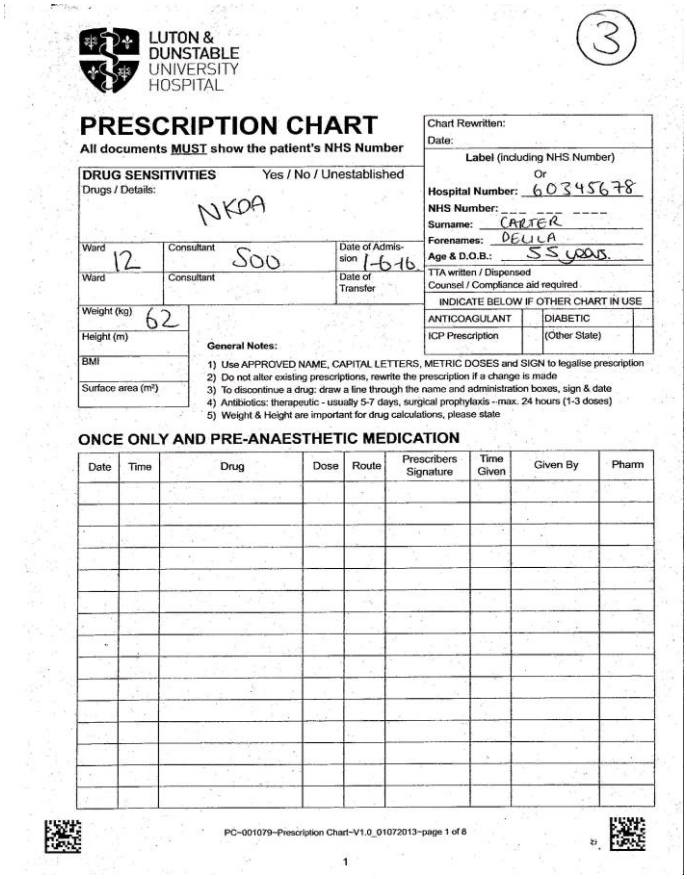
(repeated 11 times)

Date	Time	Dose	Given By	Date	Time	Dose	Given By	Date	Time	Dose	Given By

PC-001079-Prescription Chart-V1.0_01072013-page 4 of 8

4

Medication chart 3:



LUTON & DUNSTABLE UNIVERSITY HOSPITAL

3

PRESCRIPTION CHART

All documents MUST show the patient's NHS Number

Chart Rewritten:
Date:

Label (including NHS Number)
Or
Hospital Number: 6034578
NHS Number:
Surname: CARTER
Forenames: DELILA
Age & D.O.B.: 55 years.
TTA written / Dispensed
Counsel / Compliance aid required

INDICATE BELOW IF OTHER CHART IN USE

ANTICOAGULANT	DIABETIC
ICP Prescription	(Other State)

DRUG SENSITIVITIES Yes / No / Unestablished
Drugs / Details: NKDA

Ward	Consultant	Date of Admission
12	SOO	1-6-16
Ward	Consultant	Date of Transfer

Weight (kg) 62
Height (m)
BMI
Surface area (m²)

General Notes:

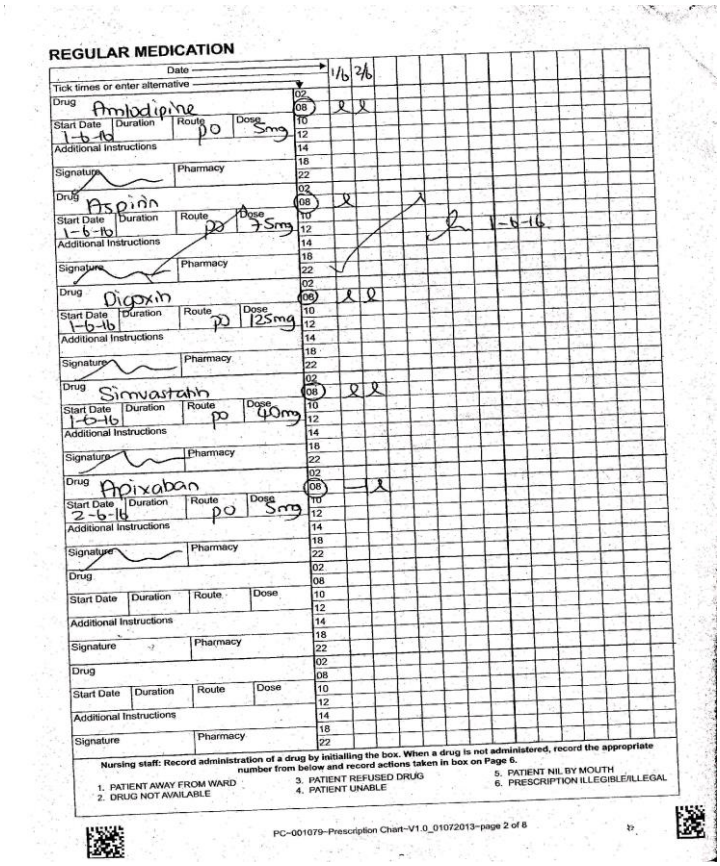
1) Use APPROVED NAME, CAPITAL LETTERS, METRIC DOSES and SIGN to legalise prescription
2) Do not alter existing prescriptions, rewrite the prescription if a change is made
3) To discontinue a drug: draw a line through the name and administration boxes, sign & date
4) Antibiotics: therapeutic - usually 5-7 days, surgical prophylaxis - max. 24 hours (1-3 doses)
5) Weight & Height are important for drug calculations, please state

ONCE ONLY AND PRE-ANAESTHETIC MEDICATION

Date	Time	Drug	Dose	Route	Prescribers Signature	Time Given	Given By	Pharm

PC-001079-Prescription Chart-V1.0_01072013-page 1 of 8

1



REGULAR MEDICATION

Date: 1/6 2/6
Tick times or enter alternative

Drug: Amlodipine
Start Date: 1-6-16 | Duration | Route: po | Dose: 5mg
Additional Instructions
Signature | Pharmacy

Drug: Aspirin
Start Date: 1-6-16 | Duration | Route: po | Dose: 75mg
Additional Instructions
Signature | Pharmacy

Drug: Digoxin
Start Date: 1-6-16 | Duration | Route: po | Dose: 125mg
Additional Instructions
Signature | Pharmacy

Drug: Simvastatin
Start Date: 1-6-16 | Duration | Route: po | Dose: 40mg
Additional Instructions
Signature | Pharmacy

Drug: Apixaban
Start Date: 2-6-16 | Duration | Route: po | Dose: 5mg
Additional Instructions
Signature | Pharmacy

Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

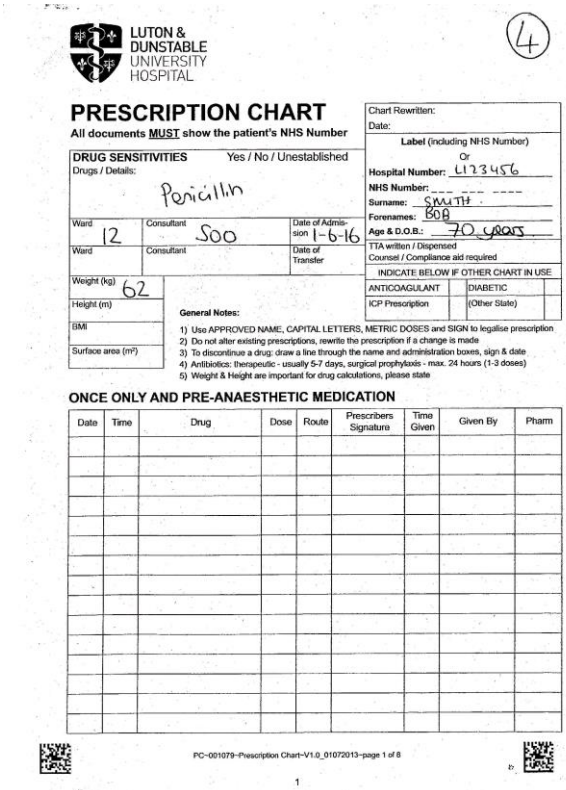
1-6-16

Nursing staff: Record administration of a drug by initialling the box. When a drug is not administered, record the appropriate number from below and record actions taken in box on Page 6.

1. PATIENT AWAY FROM WARD
2. DRUG NOT AVAILABLE
3. PATIENT REFUSED DRUG
4. PATIENT UNABLE
5. PATIENT NIL BY MOUTH
6. PRESCRIPTION ILLEGIBLE/ILLEGAL

PC-001079-Prescription Chart-V1.0_01072013-page 2 of 8

Medication chart 4:



LUTON & DUNSTABLE UNIVERSITY HOSPITAL

4

PRESCRIPTION CHART

All documents MUST show the patient's NHS Number

Chart Rewritten:
Date:
Label (including NHS Number)
Or
Hospital Number: L123456
NHS Number:
Surname: SMITH
Forenames: BOB
Age & D.O.B.: 70 years
TTA written / Dispensed
Counsel / Compliance aid required

DRUG SENSITIVITIES Yes / No / Unestablished
Drugs / Details: Penicillin

Ward	Consultant	Date of Admission
12	S00	1-6-16
Ward	Consultant	Date of Transfer

Weight (kg) 62
Height (m)
BMI
Surface area (m²)

INDICATE BELOW IF OTHER CHART IN USE

ANTICOAGULANT	DIABETIC
ICP Prescription	(Other State)

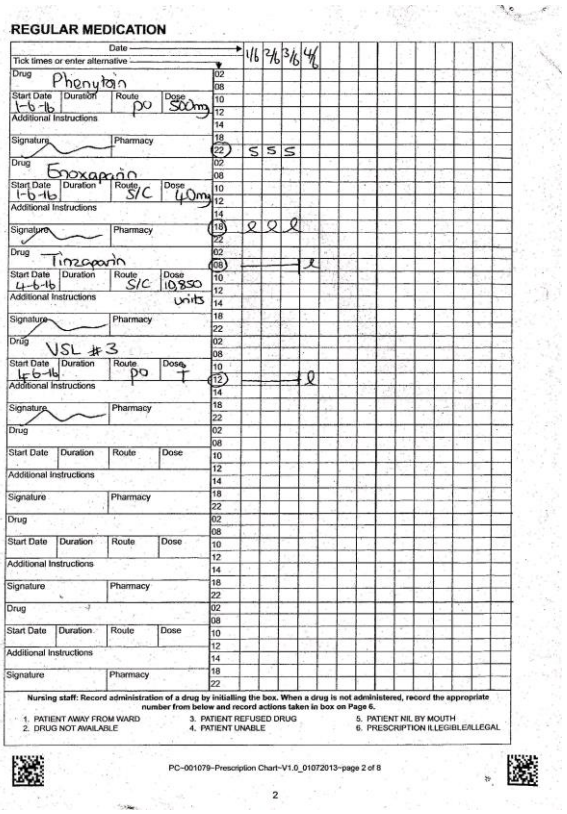
General Notes:
1) Use APPROVED NAME, CAPITAL LETTERS, METRIC DOSES and SIGN to legalise prescription
2) Do not alter existing prescriptions, rewrite the prescription if a change is made
3) To discontinue a drug: draw a line through the name and administration boxes, sign & date
4) Antibiotics: therapeutic - usually 5-7 days, surgical prophylaxis - max. 24 hours (1-3 doses)
5) Weight & Height are important for drug calculations, please state

ONCE ONLY AND PRE-ANAESTHETIC MEDICATION

Date	Time	Drug	Dose	Route	Prescribers Signature	Time Given	Given By	Pharm

PC-001079-Prescription Chart-V1.0_01072013-page 1 of 8

1



REGULAR MEDICATION

Date → 1/6 2/6 3/6 4/6
Tick times or enter alternative →

Drug: Phenytoin
Start Date: 1-6-16 | Duration | Route: po | Dose: 500mg
Additional Instructions
Signature | Pharmacy
02 08 10 12 14 18 22 02 08 10 12 14 18 22 02 08 10 12 14 18 22 02 08 10 12 14 18 22
(22) S S S

Drug: Enoxaparin
Start Date: 1-6-16 | Duration | Route: S/C | Dose: 40mg
Additional Instructions
Signature | Pharmacy
(18) ✓ ✓ ✓

Drug: Tinzaparin
Start Date: 4-6-16 | Duration | Route: S/C | Dose: 10,850 units
Additional Instructions
Signature | Pharmacy
(08) ✓

Drug: VSL #3
Start Date: 5-6-16 | Duration | Route: po | Dose: 1
Additional Instructions
Signature | Pharmacy
(12) ✓

Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

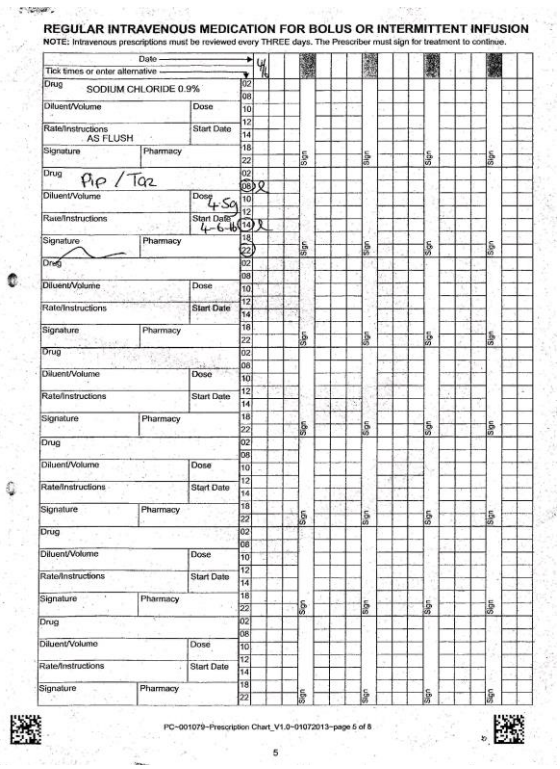
Drug
Start Date | Duration | Route | Dose
Additional Instructions
Signature | Pharmacy

Nursing staff: Record administration of a drug by initialling the box. When a drug is not administered, record the appropriate number from below and record actions taken in box on Page 6.

1. PATIENT AWAY FROM WARD
2. DRUG NOT AVAILABLE
3. PATIENT REFUSED DRUG
4. PATIENT UNABLE
5. PATIENT NIL BY MOUTH
6. PRESCRIPTION ILLEGIBLE/ILLEGAL

PC-001079-Prescription Chart-V1.0_01072013-page 2 of 8

2



REGULAR INTRAVENOUS MEDICATION FOR BOLUS OR INTERMITTENT INFUSION

NOTE: Intravenous prescriptions must be reviewed every THREE days. The Prescriber must sign for treatment to continue.

Date →
Tick times or enter alternative →

Drug: SODIUM CHLORIDE 0.9%
Diluent/Volume | Dose
Rate/Instructions: AS FLUSH | Start Date
Signature | Pharmacy

Drug: PIP / Taz
Diluent/Volume | Dose: 4.5g
Rate/Instructions | Start Date: 4-6-16
Signature | Pharmacy
02 08 (10) 12 14 18 22 02 08 10 12 14 18 22 02 08 10 12 14 18 22 02 08 10 12 14 18 22
(08) ✓ (10) ✓ (22)

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Drug
Diluent/Volume | Dose
Rate/Instructions | Start Date
Signature | Pharmacy

Sign

PC-001079-Prescription Chart_V1.0-01072013-page 5 of 8

5

Appendix A9.1: High-risk medicines included in development of the MOAT

High-risk medicines category	Medicines included
Anticoagulants / direct oral anticoagulants	Apixaban; edoxaban; dabigatran; rivaroxaban; warfarin
Therapeutic heparin	Dalteparin; enoxaparin; fondaparinux; tinzaparin
Anti-diabetic medication	Acarbose; alogliptin; glibenclamide; gliclazide; glimepiride; insulin; linagliptin; metformin; nateglinide; pioglitazone; repaglinide; saxagliptin; sitagliptin; vildagliptin
Opiates (excluding codeine, tramadol, meptazinol & dihydrocodeine)	Buprenorphine; diamorphine; fentanyl; methadone; morphine; oxycodone; tapentadol
Aminoglycosides & glycopeptides	Amikacin; gentamicin; teicoplanin; vancomycin injection
Systemic antimicrobials (excluding aminoglycosides & glycopeptides)	Aciclovir; amoxicillin; atovaquone; azithromycin; aztreonam; benzylpenicillin; caspofungin; cephalexin; cefotaxime; ceftriaxone; ciprofloxacin; clarithromycin; clindamycin; co-amoxiclav; colistimethate; co-trimoxazole; dapsone; daptomycin; demeclocycline; doxycycline; ertapenem; erythromycin; ethambutol; famciclovir; flucloxacillin; fluconazole; fosfomycin; isoniazid; itraconazole; levofloxacin; linezolid; meropenem; metronidazole; moxifloxacin; oseltamivir; penicillin V; piperacillin with tazobactam; posaconazole; riamet; rifampicin; rifater; sodium fusidate; temocillin; terbinafine; tigecycline; trimethoprim
Epilepsy medicines	Carbamazepine; gabapentin; lamotrigine; levetiracetam; phenobarbital; phenytoin; pregabalin; primidone; sodium valproate; topiramate; valproate semisodium; zonisamide; lacosamide
Antipsychotics	Amisulpride; aripiprazole; chlorpromazine; clozapine; flupentixol; haloperidol; levomepromazine; olanzapine; quetiapine; risperidone; sulpiride; trifluoperazine; zuclopenthixol
Antiarrhythmics	Adenosine; amiodarone; digoxin; dronedarone; flecainide; verapamil
Antidepressants	Amitriptyline; citalopram; clomipramine; dosulepin; doxepin; duloxetine; escitalopram; fluoxetine; mirtazapine; nortripyline; paroxetine; sertraline; trazodone; venlafaxine
Other	• Anti-retrovirals: atripla; darunavir; raltegravir; ritonavir; truvada; tenofovir • Theophylline & aminophylline • Cytotoxics: axitinib; capecitabine; erlotinib; hydroxycarbamide; methotrexate; osimertinib • Immunosuppressants (excluding corticosteroids): azathioprine; ciclosporin; leflunomide; mycophenolate mofetil; tacrolimus; adalimumab • Lithium • Medicines for Parkinson's disease: co-beneldopa; co-careldopa; orphenadrine; pramipexole; ropinirole; rotigotine; selegiline; stalevo; trihexphenidyl

Appendix A9.2: Range for candidate predictor values before and after multiple imputation

Candidate predictor	Original / imputed data	Minimum value	Maximum value
Socioeconomic status, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]	Original	0	100
	Imputed	0	100
Body mass index (kg/m^2)	Original	10.6	40.6
	Imputed	10.6	47.6
Renal function - estimated glomerular filtration rate‡ (ml/min/1.73m^2)	Original	3	161
	Imputed	3	190
Serum albumin (g/L)	Original	20	55
	Imputed	14.9	55
Serum potassium (mmol/L)	Original	2.3	5.9
	Imputed	2.3	6.3
Serum sodium (mmol/L)	Original	111	170
	Imputed	111	170
White cell count (10^9/L)	Original	0.3	20.7
	Imputed	0.3	20.9
Platelet count (10^9/L)	Original	50	490
	Imputed	50	570

‡ Glomerular filtration rate estimated using modified Modification of Diet in Renal Disease (MDRD) equation[121]

Appendix A9.3: Comparison of multivariable regression coefficients for complete-case and multiply imputed datasets (excluding body mass index)

Predictor	Complete-cases (observations = 1,440*)			Multiple imputation (observations = 1,503)		
	Regression coefficient†	Standard error	*p* value‡	Regression coefficient†	Standard error	*p* value‡
Demographic						
Age	0.0324	0.042	0.438	0.0425	0.041	0.298
Socioeconomic status	0.0430	0.021	**0.040**	0.0450	0.021	**0.029**
Patient related						
Previous allergy	0.265	0.121	**0.028**	0.260	0.118	**0.028**
Number of hospital admissions in previous 6 months	0.0254	0.076	0.738	0.0257	0.075	0.732
Primary diagnosis:						
Endocrine and metabolic	Base category		**0.035**	Base category		**0.020**
Nervous system and mental disorders	0.274	0.306		0.319	0.303	
Cardiovascular system	-0.211	0.283		-0.227	0.282	
Respiratory system	-0.354	0.281		-0.318	0.278	
Gastrointestinal system	-0.673	0.319		-0.660	0.317	
Genitourinary system	-0.085	0.302		-0.058	0.300	
Musculoskeletal-intergumentary	0.036	0.333		0.033	0.329	
All other categories	-0.083	0.289		-0.0085	0.286	
Number of comorbidities	0.137	0.033	**<0.001**	0.137	0.033	**<0.001**
History of dementia	-0.226	0.196261	0.249	-0.259	0.193	0.181
Medicines related						
Number of medicines	0.0239	0.018	0.185	0.0230	0.018	0.195
Use of high-risk medicines:						
Anticoagulants	0.102	0.156	0.516	0.106	0.153	0.489
Therapeutic heparin	0.279	0.181	**0.122**	0.266	0.179	**0.137**
Anti-diabetic medication	0.151	0.154	0.329	0.217	0.151	**0.151**
Opiates	0.0414	0.198	0.834	0.0138	0.197	0.944
Aminoglycosides and glycopeptides	0.400	0.229	**0.081**	0.330	0.225	**0.143**
Other antimicrobials	0.334	0.150	**0.026**	0.361	0.147	**0.014**
Epilepsy medicines	0.445	0.167	**0.008**	0.478	0.165	**0.004**
Antipsychotics	0.131	0.238	0.583	0.164	0.236	0.487
Antiarrhythmics	-0.0809	0.204	0.692	-0.0578	0.201	0.773

Continued from previous page…

Predictor	Complete-cases (observations = 1,440*)			Multiple imputation (observations = 1,503)		
	Regression coefficient†	Standard error	*p* value‡	Regression coefficient†	Standard error	*p* value‡
Antidepressants	0.209	0.139	**0.134**	0.203	0.138	**0.140**
Other high-risk medicines	0.139	0.218	0.525	0.123	0.215	0.566
Parenteral administration	-0.0194	0.152	0.899	0.0408	0.149	0.784
Laboratory results						
Estimated glomerular filtration rate	-0.0417	0.020	**0.038**	-0.0358	0.020	**0.071**
Liver disease	-0.0709	0.195	0.717	-0.0962	0.194	0.620
Serum albumin	0.00036	0.012	0.975	0.00051	0.012	0.965
Serum potassium	-0.123	0.102	0.226	-0.128	0.100	0.204
Serum sodium	-0.0079	0.011	0.490	-0.0063	0.011	0.575
White cell count	0.0262	0.015	**0.086**	0.0225	0.015	**0.136**
Platelet count	0.0019	0.007	0.781	0.0029	0.007	0.665

* Following inclusion of values determined using common sense solutions[182]

† Relationship between the independent and dependent variable (amount of increase in predicted log odds of the outcome event that would be predicted by a one unit increase in the independent variable)

‡ Values with a significance level of $p<0.157$ shown in bold

Appendix A9.4: Univariable and multivariable association between predictors and outcome events

Predictor	Univariable analysis		Multivariable analysis	
	Odds ratio† (95% CI)	*p* value‡	Odds ratio† (95% CI)	*p* value‡
Demographic				
Age/10 (years)	1.18 (1.12 to 1.25)	<0.001	1.04 (0.96 to 1.13)	0.352
Socioeconomic status/10, ranked using English Indices of Deprivation 2015 – Index of Multiple Deprivation[149]*	1.04 (1.00 to 1.08)	0.044	1.05 (1.00 to 1.09)	0.034
Patient related				
Previous allergy	1.67 (1.35 to 2.06)	<0.001	1.30 (1.02 to 1.67)	0.035
Body mass index (kg/m^2)	1.01 (0.99 to 1.03)	0.610	0.100 (0.97 to 1.02)	0.810
Number of hospital admissions in previous 6 months	1.19 (1.05 to 1.35)	0.008	1.02 (0.88 to 1.191)	0.757
Primary diagnosis:				
Endocrine and metabolic	1.34 (0.81 to 2.2)	0.260	1.01 (0.57 to 1.78)	0.977
Nervous system and mental disorders	1.40 (0.92 to 2.1)	0.113	1.40 (0.88 to 2.21)	0.158
Cardiovascular system	1.19 (0.84 to 1.67)	0.329	0.80 (0.54 to 1.20)	0.279
Respiratory system	1.17 (0.83 to 1.65)	0.359	0.73 (0.48 to 1.10)	0.137
Gastrointestinal system	0.64 (0.41 to 0.99)	0.048	0.52 (0.31 to 0.87)	0.013
Genitourinary system	1.53 (1.01 to 2.33)	0.046	0.95 (0.59 to 1.55)	0.847
Musculoskeletal-intergumentary systems	1.67 (1.03 to 2.71)	0.036	1.05 (0.61 to 1.81)	0.859
All other categories	Base category		Base category	
Number of comorbidities	1.24 (1.18 to 1.30)	<0.001	1.15 (1.07 to 1.24)	<0.001
History of dementia	1.05 (0.75 to 1.46)	0.779	0.76 (0.52 to 1.14)	0.182
Laboratory results				
Estimated glomerular filtration rate/10* (ml/min/1.73m^2)	0.92 (0.89 to 0.95)	<0.001	0.96 (0.93 to 1.00)	0.076
Liver disease	0.90 (0.65 to 1.26)	0.548	0.90 (0.61 to 1.33)	0.609
Serum albumin (g/L)	0.97 (0.95 to 0.99)	<0.001	1.00 (0.98 to 1.02)	0.958
Serum potassium (mmol/L)	1.09 (0.91 to 1.29)	0.358	0.88 (0.72 to 1.08)	0.213
Serum sodium (mmol/L)	0.99 (0.97 to 1.01)	0.257	0.99 (0.97 to 1.02)	0.583
White cell count (10^9/L)	1.03 (1.00 to 1.05)	0.021	1.02 (0.99 to 1.06)	0.143

Continued from previous page…

Predictor	Univariable analysis		Multivariable analysis	
	Odds ratio† (95% CI)	***p* value‡**	**Odds ratio† (95% CI)**	***p* value‡**
Platelet count/10 (10^9/L)	1.00 (0.99 to 1.01)	0.682	1.00 (0.99 to 1.02)	0.677
Medicines related				
Number of medicines	1.10 (1.07 to 1.13)	<0.001	1.02 (0.99 to 1.06)	0.194
Use of high-risk medicines:				
Anticoagulants	1.50 (1.16 to 1.92)	0.002	1.12 (0.82 to 1.52)	0.479
Therapeutic heparin	1.32 (0.99 to 1.76)	0.057	1.31 (0.92 to 1.89)	0.139
Anti-diabetic medication	1.71 (1.33 to 2.21)	<0.001	1.25 (0.92 to 1.70)	0.151
Opiates	1.37 (0.97 to 1.93)	0.073	1.01 (0.69 to 1.50)	0.951
Aminoglycosides and glycopeptides	1.74 (1.17 to 2.59)	0.006	1.40 (0.89 to 2.20)	0.147
Other antimicrobials	1.71 (1.38 to 2.13)	<0.001	1.44 (1.07 to 1.95)	0.018
Epilepsy medicines	1.92 (1.44 to 2.55)	<0.001	1.63 (1.14 to 2.31)	0.007
Antipsychotics	1.28 (0.84 to 1.95)	0.255	1.18 (0.74 to 1.89)	0.492
Antiarrhythmics	1.40 (1.00 to 1.96)	0.053	0.94 (0.63 to 1.41)	0.775
Antidepressants	1.50 (1.18 to 1.90)	0.001	1.23 (0.93 to 1.63)	0.148
Other high-risk medicines	1.33 (0.91 to 1.94)	0.143	1.13 (0.73 to 1.73)	0.585
Parenteral administration	1.46 (1.17 to 1.83)	<0.001	1.04 (0.77 to 1.40)	0.790

* Deprivation rank based on patients' postcode, shown as the ranked position as a percentage of all neighbourhoods in England (where 1 is the most deprived)

† Measure of association between exposure and outcome event (the odds that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of exposure)

‡ Test for difference between admissions with and without occurrence of outcome event. Obtained from regression modelling

Age, socioeconomic status, estimated glomerular filtration rate, and platelet count analysed as deciles to aid interpretability (i.e. after dividing the actual value by ten)

Appendix A10.1: Participant information sheet (MOAT assessment)

LUTON & DUNSTABLE UNIVERSITY HOSPITAL

West Hertfordshire Hospitals NHS

NHS Trust

Development of the Medicines Optimisation Assessment Tool (MOAT)

Targeting hospital pharmacists' input to reduce risks and improve patient outcomes

Participant information sheet

We'd like to invite you to take part in our research study. Joining the study is entirely up to you, before you decide we would like you to understand why the research is being done and what it would involve for you. We will go through this information sheet with you, to help you decide whether or not you would like to take part and answer any questions you may have. We'd suggest this should take about 5 minutes. Please feel free to talk to others about the study if you wish.

THANK YOU FOR READING THIS LEAFLET

UCL SCHOOL OF PHARMACY

UCL

Contact for further information

If you would like any further information on any aspect of the study, then do not hesitate to contact us.

Cathy Geeson
Clinical Research Fellow and Pharmacist
Phone number: 01582 497168
Email: cathy.geeson@ldh.nhs.uk

Version 1.0: 16 September 2017
IRAS ID: 197298

Please read the following information and ask us if there is anything that is not clear or if you would like more information. All proposals for research involving human subjects are reviewed by an ethics committee before they can proceed. This study has been approved by the Proportionate Review Service Sub-Committee of the National Health Service (NHS) Research Ethics Committee Wales REC 7 (16/WA/0016), and Health Research Authority (project ID 197298).

Who is organising this study?

The research is being carried out by Luton and Dunstable University Hospital NHS Foundation Trust, West Hertfordshire Hospitals NHS Trust, and University College London. It has been funded by the National Institute for Health Research (funded by the NHS to improve the health and wealth of the nation through research).

What is the purpose of this study?

Our overall aim is to develop a prediction-tool, the Medicines Optimisation Tool or MOAT, which can be used by pharmacists to identify patients at highest risk of medication related problems while in hospital. This would permit pharmacy services to be targeted appropriately.

How was the MOAT developed?

We have collected information from 1,500 patients admitted to medical wards at the Luton and Dunstable Hospital and Watford General Hospital during 2016. Pharmacists collected information on medication related problems experienced by study patients as part of their routine daily clinical assessments, and information was also collected by the research team on potential risk factors (known as predictors). These included age, use of high-risk medicines and laboratory results. These data were then analysed to find out which predictors were associated with medication related problems, and a scoring system developed.

What further research is needed?

We would now like to find out if practising pharmacists consider the MOAT to be clinically sensible. This includes if:

- any obvious predictors are missing;
- the method of grouping the predictors is reasonable;
- the predictors seem appropriate;
- time taken to use the MOAT is reasonable.

What would taking part involve?

We would like you to consider volunteering to assess the MOAT. There are two parts to the assessment, and you can volunteer to take part in one or both. First we plan to use a consensus development technique to generate agreement from nine pharmacists on how clinically sensible the MOAT is. This will involve attending two meetings, each lasting approximately two hours. After the first meeting responses will be summarised and redistributed for discussion at the next meeting. We also plan to calculate the average time taken to apply the MOAT. For this we need four volunteers, who will each use the MOAT to calculate a risk score for five patients (without acting on the findings). We estimate this should take each volunteer approximately one hour.

Participation is voluntary, and deciding not to participate will not affect you in any way. You may also withdraw at any time if you change your mind without giving a reason.

Is the study confidential?

Yes. All information collected will be kept strictly confidential. We will not include names, or identifying information in any reports.

Concerns or complaints?

If you wish to complain, or have any concerns about any aspect of this study then you should speak to the researcher.

Thank you for your time
Please ask us if you would like more information

Appendix A10.2: Participant consent form (MOAT assessment)

LUTON & DUNSTABLE UNIVERSITY HOSPITAL
CLINICAL EXCELLENCE, QUALITY & SAFETY
Lewsey Road Luton LU4 0DZ
Tel: 01582 49 11 66 www.ldh.nhs.uk

IRAS ID: 197298

CONSENT FORM

Title of Project: Development of the Medicines Optimisation Assessment Tool (MOAT) - Targeting hospital pharmacists' input to reduce risks and improve patient outcomes

Name of Researcher: Cathy Geeson

Please initial box

1. I confirm that I have read the information sheet dated 16 September 2017 (version1.0) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason. ☐
3. I understand that my responses will be used to guide future development of the Medicines Optimisation Assessment Tool (MOAT), and may be shared anonymously with other researchers. ☐
4. I agree to take part in the first part of the assessment of the MOAT for the above study (the consensus meetings to discuss clinical sensibility). ☐
5. I agree to take part in the second part of the assessment of the MOAT for the above study (to calculate the average time taken to use the MOAT). ☐

Name of Participant	Date	Signature
Name of Person taking consent	Date	Signature

Version 1.0: 16 September 2017

Appendix A10.3: Consensus score sheet (MOAT assessment)

Assessment of Medicines Optimisation Assessment Tool (MOAT) - consensus score sheet

Name of participant:……..

- Listed below are five statements. Please review each statement and indicate the extent to which you agree or disagree. A score of one indicates total disagreement, and nine indicates total agreement
- A nine-point scale has been used to permit responses to be categorised, with scores of 1-3 representing disagreement, scores of 4-6 representing an equivocal response, and scores of 7-9 representing agreement
- Please circle a score, and give brief reason(s) for your score in the space provided

1. The choice of risk factors is appropriate.

Disagreement			Equivocal			Agree		
1	2	3	4	5	6	7	8	9

Reason(s) for score:

2. The presentation of the MOAT is reasonable.

Disagreement			Equivocal			Agree		
1	2	3	4	5	6	7	8	9

Reason(s) for score:

3. The MOAT has the potential to be ‘usable’ in clinical practice:

3a. The MOAT is simple to interpret.

Disagreement			Equivocal			Agree		
1	2	3	4	5	6	7	8	9

Reason(s) for score:

3b. The time it takes to use the MOAT is reasonable.

Disagreement			Equivocal			Agree		
1	2	3	4	5	6	7	8	9

Reason(s) for score:

4. The proposed ‘decision thresholds’ for the creation of risk groups (as discussed during the meeting) are appropriate.

Disagreement			Equivocal			Agree		
1	2	3	4	5	6	7	8	9

Reason(s) for score:

Thank-you for your participation

Appendix A10.4: Workload implication (MOAT assessment)

Assessment of Medicines Optimisation Assessment Tool (MOAT) – time taken to apply the MOAT

Name of participant:……..

- Please use the MOAT to calculate a risk score (probability of experiencing a moderate or severe preventable medication related problem) for 5 patients
- Ideally these patients should have been recently admitted to hospital (i.e. all data should be unknown to you prior to the assessment)
- Record the actual time taken to complete the task, that is, please do not count travel time to or from wards, or distractions experienced (e.g. responding to queries from ward staff)
- Ideally, use a stopwatch, which can be paused if necessary, and record the time taken (minutes and seconds)
- Where data collection is performed for all patients together, (e.g. ICE results, and JAC information) please record the total time, then split as appropriate)
- It is not necessary to record patient identifying details

Please note, you will need to obtain the following data:

Number of comorbidities	A count of the number of comorbidities recorded
Number of 'regular' medicines prescribed	All regular medicine on the first full day of admission **Exclude** 'when required' and 'once only' medicines, dietary products, non-medicated topical products (e.g. emollients), wound dressings, oxygen
Age	At admission
Serum creatinine	First result following admission
Gender	Male or female
Ethnicity	Black or non-black (used to calculate eGFR)
White cell count	First result following admission
Documented allergy	YES/NO
Whether patient treated with systemic aminoglycosides or glycopeptides	i.e. gentamicin or vancomycin YES/NO
Whether patient treated with other systemic antimicrobial(s)	i.e. other than aminoglycosides & glycopeptides YES/NO
Whether patient treated with 'epilepsy medicines'	YES/NO
Primary diagnosis	i.e. Whether admission related to: • Nervous system or mental disorder • Respiratory system • Gastrointestinal system

Patient	Time taken (minutes and seconds)
1	
2	
3	
4	
5	

ICE = Sunquest Integrated Clinical Environment™ (laboratory data electronic reporting system), JAC = electronic prescribing record system, eGFR = estimated glomerular filtration rate

Appendix A10.5: Written comments from consensus meetings

Statement 1 – The choice of risk factors is appropriate

Comments (meeting 1):

- I agree with the risk factors, as more we add, the more time it would take to complete & may put people off.
- Understood rationale for risk factors, however would like anticoagulants included.
- If the data can be pulled across electronically then more risk factors should be added e.g. socioeconomic.
- Having considered the options, I still feel anticoagulants & insulins should be on the drug section.
- Lacking risk factors (e.g. anticoagulants) will be captured by co-morbidities data.
- Discussed the reason for not including high-risk drugs / conditions that are not a significant representation of patients → the focus is on co-morbidities which if noted properly would indicate associated drug groups to be used.
- Those risk factors I thought would be in there have been shown not to significantly increase risk.

Comments (meeting 2):

- There is room to add more if required, since it's electronic it will be very easy to amend data.
- The information that anticoagulants etc do not make a significant difference should be communicated so we feel confident in using the system without these categories.
- I'm still of the opinion that insulin and anticoagulants should be added even though I know I will make little difference to the scoring.
- Increased score to 8 due to capability of the tool to include anticoagulants / anti-diabetics if deemed appropriate.
- I have not changed my score as although other risk factors have been found not to be significant I feel adding anti-diabetics / anticoagulants would make others "trust" the score more + feel happier about using this tool.
- OK as it was if explain why some risk factors excluded before people use tool, need to draw line somewhere. Don't want it to be too onerous to use to start with (may put people off).

Statement 2 – The presentation of the MOAT is reasonable

Comments (meeting 1):

- All the information is on one screen.
- Generally, I really like the presentation. Appears to be easy to use & easy to read, Needs a few minor adjustments as discussed within the group. I like the green colour scheme.
- Keep the box (MOAT scoring screen) on top of other screens when clicking on other screens.
- Like the information buttons. Some concern over 'race' button (i.e. may not be able to establish race by 'looking' at a patient). All on one screen. Put in ability to increase size of screen. Colour – I like it! 'Units' for creatinine – able to choose

option (i.e. add 'mg/dl' as an alternative option). Get rid of Excel background. Colour blindness!

- Easy to use. Clear & legible. Remove Excel background. Simplified → colour scoring system + percentage score (good visuals). Hassle free. Restricted options allowing for ease of use without complicating matters. Room for improvement: submit button → can this calculate eGFR + risk automatically; risk category writing → BOLDER + BIGGER.
- "Green" low score – how will it show up on green background.

Comments (meeting 2):

- Higher resolution and size. Tool will grow with lots more data or function buttons.
- After using the MOAT tool I found the presentation really easy to use. All changes we have discussed are easily made.

Statement 3a – The MOAT is simple to interpret

Comments (meeting 1):

- Anyone should be able to use, provided they have the right training / information.
- Really easy to follow. Jess made a good point regarding explanation of "Probability %".
- Very easy to interpret. Easy to use. Give ranges for risk scoring – explanation buttons.
- Intention is fantastic to save time focus on those most high-risk / vulnerable patient → clear colour system → + % score to further rationalise which patients to see. Can further improve by adding a key → explaining the scoring system / categories (how high / medium / low-risk).
- Statement for what % is actually showing, Key for what high / med / low means + what the cut offs are.

Comments (meeting 2):

- No change to score

Statement 3b – The time it takes to use the MOAT is reasonable

Comments (meeting 1):

- Obtaining information electronically should not take long, via patient or notes may take a bit longer.
- Difficult to answer until it's been tested on the wards. Think the benefits will outweigh the time taken as it will save time later → help prioritise high-risk patients.
- Not sure, have to complete bedplans, pharmacy friend, board round. If this tool gets added on it may slow our work rate down?
- This is the part that causes me most concern.
- Reasonable for an area with lower levels of new patients, may not be in acute setting.
- Theoretical time taken to complete the tool would be beneficial as it would allow you to spend your time more wisely on the ward to high-risk groups. In practise → too time consuming on acute e.g. 21 new patients daily???
- Unsure until put into practise. May be more useful on one ward than another.

Comments (meeting 2):

- As before, should not take long provide you have access to all the info.
- 3 mins using tool beats 20 mins medicines reconciliation.
- Changed score after using – very quick i.e. only if don't need to use patients notes.

Statement 4 – The proposed 'decision thresholds' for the creation of risk groups (as discussed during the meeting) are appropriate

Comments (meeting 1):

- Current gap appropriate, more can be added in future.
- Low-risk should still be seen during their inpatient stay. Understood rationale behind risk groups.
- I feel the % are more important but the thresholds do help & allow flexibility.
- The 'have to see' is useful. I like the % scoring – I can make my decision based on this.
- Good prioritisation tool + scoring system to further allow you to focus on patients within categories.
- Need to try out to see how this works in practise – will be saturated with "Red" patients. Able to use % rather than categories to re-prioritise. Don't wish to be in 'acute' (the acute assessment wards) when all low-risk patients are not seen at all by pharmacy.

Comments (meeting 2):

- Still think people will use percentage risk over groups.

Appendix A10.6: Moderate or severe preventable medication related problems (MSP MRPs) experienced by 'low-risk' patients

Study code	Predicted risk probability (%)	Details of moderate or severe preventable medication related problems (MSP MRPs) experienced
W0104	12.2	Subtherapeutic dose of prophylactic enoxaparin (based on weight)
W0374	13.1	No prescription chart written - therefore patient's regular felodipine not prescribed
		No prescription chart written - therefore patient's regular morphine liquid (10mg/5ml) not prescribed
		No prescription chart written - therefore patient's regular omeprazole not prescribed
		Patient recently prescribed apixaban, advised to review whether aspirin still indicated
W0353	14.6	Medicines reconciliation discrepancy - ramipril not prescribed
W0126	14.9	Prescribed full dose paracetamol (patient jaundiced & liver function impaired)
W0322	15.1	Venous thromboembolism (VTE) assessment not completed - unclear if prophylaxis required
M0948	15.2	Medicines reconciliation discrepancy - beclometasone inhaler not prescribed
W0451	15.9	Medicines reconciliation discrepancy - oral contraceptive pill not prescribed
		Co-amoxiclav 1.2g three times daily prescribed orally - had intended parenteral
M0914	16.0	Dexamethasone prescribed as 8mg twice daily for 14 days only - had intended to prescribe as a reducing regimen
M0609	16.4	Medicines reconciliation discrepancy - ivabradine not prescribed
		Medicines reconciliation discrepancy - clopidogrel not prescribed
M0344	16.6	Medicines reconciliation discrepancy - olanzapine not prescribed
M0513	17.1	Medicines reconciliation discrepancy - bisoprolol dose too low
W0445	17.3	Medicines reconciliation discrepancy - levothyroxine not prescribed
M0299	17.5	Medicines reconciliation discrepancy - fluticasone inhaler not prescribed
M1098	17.8	Medicines reconciliation discrepancy - bisoprolol not prescribed
M0987	18.0	Medicines reconciliation discrepancy - sertraline not prescribed
		Medicines reconciliation discrepancy - morphine sulphate modified release tablets not prescribed
		Medicines reconciliation discrepancy - morphine liquid (10mg/5ml) not prescribed
		Medicines reconciliation discrepancy - gabapentin not prescribed
		Medicines reconciliation discrepancy - omeprazole not prescribed (patient also takes naproxen)
M0577	18.0	Incorrect rate of aminophylline infusion prescribed (40 mg/hour instead of 20mg/hour)
M0004	18.3	Medicines reconciliation discrepancy - mirtazapine not prescribed
W0101	18.8	Medicines reconciliation discrepancy - citalopram not prescribed
		Subtherapeutic dose of prophylactic enoxaparin (based on weight)
W0299	18.9	Warfarin prescribed but administration time not specified
W0303	19.2	Medicines reconciliation discrepancy - amlodipine not prescribed
M0227	19.3	Patient prescribed sliding scale insulin - regular long-acting insulin not prescribed
W0177	19.4	Medicines reconciliation discrepancy - quetiapine not prescribed

Continued from previous page…

Study code	Predicted risk probability (%)	Details of MSP MRP
M0709	19.6	Prophylactic penicillin V not prescribed on discharge prescription (to restart after current course of co-amoxiclav)
M0017	19.6	Fondaparinux not discontinued (acute coronary syndrome diagnosis excluded)
M0812	19.7	Prescribed amlodipine on discharge prescription, not prescribed an inpatient (was wrong patient)
M0238	20.2	Reason for admission recurrent eye infection - no treatment initiated
W0337	20.3	VTE assessment completed - patient high-risk but no prophylactic heparin prescribed
M0230	20.6	Medicines reconciliation discrepancy - sertraline not prescribed
		Medicines reconciliation discrepancy - aspirin not prescribed
		Medicines reconciliation discrepancy - lansoprazole not prescribed (patient on high-dose prednisolone)
W0439	20.9	Chlordiazepoxide prescribed - no plan for reducing dose over weekend
M0983	21.3	Provisional diagnosis of stroke - only 1 antiplatelet prescribed
		Medicines reconciliation discrepancy - lansoprazole not prescribed (patient has gastro-oesphageal reflux disease)
W0264	21.4	VTE assessment completed - patient high-risk but no prophylactic heparin prescribed
M0894	21.4	No maximum frequency prescribed for 'as required' midazolam
M0690	21.7	No stop date specified for filgrastim
M0453	22.0	Medicines reconciliation discrepancy - amlodipine not prescribed
		Medicines reconciliation discrepancy - mirtazapine not prescribed
M0942	22.0	Medicines reconciliation discrepancy - ramipril not prescribed
		Medicines reconciliation discrepancy - metformin not prescribed
M0587	22.1	Medicines reconciliation discrepancy - trimethoprim not prescribed
W0369	22.2	Medicines reconciliation discrepancy - flupentixol not prescribed
M0835	22.3	Medicines reconciliation discrepancy - amlodipine not prescribed
		Medicines reconciliation discrepancy - resource thickener not prescribed
W0391	22.4	VTE assessment completed - patient high-risk but no prophylactic heparin prescribed
M0825	22.5	Medicines reconciliation discrepancy - citalopram not prescribed
		Medicines reconciliation discrepancy - memantine not prescribed
W0276	22.5	Paracetamol prescribed as 'regular' and 'as required'
M0416	22.5	Sando K not discontinued despite potassium level now above standard reference range
M0715	22.8	Meropenem prescribed for patient with documented meropenem allergy. Appears allergy status incorrect - therefore corrected
W0297	22.8	VTE assessment not completed - unclear if prophylaxis required
W0075	22.9	Morphine liquid (10mg/5ml) prescription not legible (poor handwriting and units not stated)

Continued from previous page…

Patient reference	Predicted risk probability (%)	Details of MSP MRP
M0355	23.1	Medicines reconciliation discrepancy - sertraline not prescribed
		Medicines reconciliation discrepancy - quetiapine not prescribed
		Sub-therapeutic dose of prophylactic heparin (not increased following improvement in renal function)
W0316	23.1	Medicines reconciliation discrepancy - pilocarpine eye drops not prescribed
		Medicines reconciliation discrepancy - Azopt eye drops not prescribed
		Medicines reconciliation discrepancy - Lumigan eye drops not prescribed
		Medicines reconciliation discrepancy - finasteride not prescribed
M0516	23.2	Medicines reconciliation discrepancy - latanoprost/timolol eye drops not prescribed
M0132	23.3	Allergy status not recorded
M0698	23.5	Originally on clopidogrel - admitted with 'non-ST-elevation myocardial infarction'. Could be clopidogrel non-responder - switch to ticagrelor
M0030	23.7	Prescribed ciprofloxacin for 'spontaneous bacterial peritonitis' prophylaxis (long term) - suspended during admission whilst on alternative antibiotic treatment but not re-prescribed on discharge prescription
W0135	23.7	Medicines reconciliation discrepancy - enalapril not prescribed
		Prophylactic enoxaparin not prescribed
W0437	23.8	Dose of prophylactic enoxaparin too high (based on weight)
M0845	24.0	Teicoplanin dose too high (based on weight)
W0350	24.1	Medicines reconciliation discrepancy - dose of doxazosin too low
		Medicines reconciliation discrepancy - ramipril not prescribed
M0336	24.4	Medicines reconciliation discrepancy - amlodipine not prescribed
		Medicines reconciliation discrepancy - losartan not prescribed
M0055	24.5	Medicines reconciliation discrepancy - letrozole dose too high
W0418	24.8	Patient transferred wards, medication (quetiapine) not transferred with patient
W0040	24.9	Patient's weight recorded incorrectly so prophylactic heparin dose prescribed incorrectly
M0309	24.9	Filgrastim dose too low (prescribed 250mcg daily, usual dose 300mcg OD)
M0769	25.2	Medicines reconciliation discrepancy - bimatoprost eye drops not prescribed
		Medicines reconciliation discrepancy - salbutamol inhaler not prescribed
		Patient has 'chronic pulmonary obstructive disease' - but only treatment prescribed is salbutamol - review required

Publications & conference presentations

1. Study protocol, published in BMJ Open, June 2017

Reference: Geeson C, Wei L, Franklin BD. Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study. *BMJ Open* 2017;7(6) doi: 10.1136/bmjopen-2017-017509

Open Access Protocol

BMJ Open

Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study

Cathy Geeson,[1,2] Li Wei,[2] Bryony Dean Franklin[2,3]

To cite: Geeson C, Wei L, Franklin BD. Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study. *BMJ Open* 2017;**7**:e017509. doi:10.1136/bmjopen-2017-017509

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-017509).

Received 26 April 2017
Revised 10 May 2017
Accepted 10 May 2017

[1]Pharmacy, Luton and Dunstable University Hospital, Luton, Bedfordshire, UK
[2]UCL School of Pharmacy, London, UK
[3]Centre for Medication Safety and Service Quality, Imperial College Healthcare NHS Trust, London, UK

Correspondence to
Cathy Geeson; cathy.geeson@ldh.nhs.uk

ABSTRACT

Introduction Medicines optimisation is a key role for hospital pharmacists, but with ever-increasing demands on services there is a need to increase efficiency while maintaining patient safety. The aim of this study is to develop a prognostic model, the Medicines Optimisation Assessment Tool (MOAT), which can be used to target patients most in need of pharmacists' input while in hospital.

Methods and analysis The MOAT will be developed following recommendations of the Prognosis Research Strategy partnership. Using a cohort study we will prospectively include 1500 adult patients from the medical wards of two UK hospitals. Data on medication-related problems (MRPs) experienced by study patients will be collected by pharmacists at the study sites as part of their routine daily clinical assessment of patients. Data on potential risk factors such as polypharmacy, renal impairment and the use of 'high risk' medicines will be collected retrospectively from the information departments at the study sites, laboratory reporting systems and patient medical records. Multivariable logistic regression models will then be used to determine the relationship between potential risk factors and the study outcome of preventable MRPs that are at least moderate in severity. Bootstrapping will be used to adjust the MOAT for optimism, and predictive performance will be assessed using calibration and discrimination. A simplified scoring system will also be developed, which will be assessed for sensitivity and specificity.

Ethics and dissemination This study has been approved by the Proportionate Review Service Sub-Committee of the National Health Service Research Ethics Committee Wales REC 7 (16/WA/0016) and the Health Research Authority (project ID 197298). We plan to disseminate the results via presentations at relevant patient/public, professional, academic and scientific meetings and conferences, and will submit findings for publication in peer-reviewed journals.

Trial registration number NCT02582463.

Strengths and limitations of this study

- The Medicines Optimisation Assessment Tool (MOAT) will be the first evidence-based prognostic model to identify hospitalised patients at risk of moderate or severe medication-related problems in order to permit targeting by pharmacists.
- The study will include adult patients of all ages admitted to all types of medical wards (general, emergency and elderly medicine), so will be representative of patients routinely admitted to hospital medical wards.
- The method and analysis plan are based on the Prognosis Research Strategy, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) and CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) recommendations for prognostic research.
- The study is observational, which may be subject to reporting bias and missing data.
- Only two hospitals will be included in the study; further validation, impact and implementation studies will be needed to determine whether the MOAT could be successfully employed in new settings.

INTRODUCTION

Medicines play a crucial role in maintaining health and are the most common intervention in healthcare. However, in the UK, as elsewhere, there is a growing body of evidence that there is a need to improve medicines use.[1–6] This includes the Francis and Berwick reports,[1 2] which call for a number of actions to improve patient safety and reduce avoidable harm.

Historically, adverse drug events have been the focus of studies of medication-related harm,[7] but problems can also result from suboptimal medicines use, such as ineffective treatments or subtherapeutic doses. It is estimated that only 4%–21% of patients in primary care receive optimum benefit from their medicines,[8] and it has been suggested that research efforts should also identify patients with unrealised benefits.[9] A term that

encompasses both aspects is medication-related problems (MRPs), defined as all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome.[7 10–12] This also shifts the focus from 'medication-related harm' to 'medicines optimisation', which can be described as the safe and effective use of medicines to enable the best possible outcomes.[5] Medicines optimisation is high on the English national agenda, with guidance issued by the Royal Pharmaceutical Society and the National Institute for Health and Care Excellence.[4 5]

Medicines optimisation is a key role for pharmacists,[13–16] and a number of systematic reviews conclude that addition of clinical pharmacy services to the care of hospital inpatients improves quality, safety and efficiency of patient care.[10 17 18] Ideally, pharmacists would see every patient daily, but medicines optimisation is not the only goal for hospital pharmacy services in England.[13 19] Other service developments are required, such as delivery of 7-day services[20] and the Hospital Pharmacy Transformation Programme, as set out in the recent review by Lord Carter on improving productivity and performance in English National Health Service (NHS) acute hospitals.[21] Owing to financial challenges that face the NHS, these developments often have to be achieved within existing funding through increased efficiency and innovation.[5 22 23] There have therefore been calls from international government organisations and professional bodies for effective ways for pharmacy services to target patients most in need.[15 24–28]

Clinical prioritisation has been proposed as a way to permit pharmacy services to focus on the greatest need and where clinical pharmacy input is likely to have greatest impact. This requires a method to triage patients to assign 'pharmaceutical acuity'.[28 29] There are recognised risk factors for MRPs, for example polypharmacy, renal impairment and the use of 'high risk' medicines,[30] but to target patients appropriately pharmacists need to be able to apply this knowledge effectively and consistently within their routine clinical practice.

Predicting clinical risk is well established in medicine. Tools such as cardiac-risk calculators and the Waterlow score (to assess the risk of pressure ulcers) are both used daily across the NHS.[31 32] Prediction tools to identify hospitalised patients at risk of adverse medication-related outcomes have been developed,[33–41] but the majority identify patients at risk of adverse drug reactions,[34 35] adverse drug events[36] or medication errors,[37] rather than MRPs, or are based on 'expert opinion' rather than statistical determination.[38–41]

Interest in prediction research (also known as prognosis research) has developed rapidly in recent years. It involves use of statistical methods to predict future health outcomes among people with a given baseline health status, and therefore has potential to inform clinical decision making, improve patient care and make healthcare more efficient.[42 43] Prognostic modelling is one component of prognosis research, in which multiple risk (prognostic) factors are statistically combined to predict future clinical risk for an individual patient.[44] However, many published prognostic model studies have been criticised in terms of methodological shortcomings, limiting their reliability and applicability,[44 45] as well as poor reporting, which limits the ability to effectively assess the risk of bias.[43 46] Both problems ultimately limit the usefulness of the prognostic models. The perceived inadequacies in prognostic model research prompted the recent publication of recommendations for prognosis research by the Prognosis Research Strategy (PROGRESS) partnership,[42 44 47 48] together with specific guidelines for reporting[43 46] and critically appraising[49] prognostic model research.

This study aims to address a current gap in the evidence base: the development of a methodologically sound prognostic model to target hospital patients most in need of pharmacists' input based on their risk of MRPs. The purpose of publishing this study protocol is to expand on the details already publicly available[50] to clearly specify a priori the outcome measure, prognostic factors and analysis plan. This is intended to protect against both data-driven model development (associated with overoptimistic model performance) and selective reporting. The study method is informed by the PROGRESS guidelines, and the protocol includes the key elements proposed by Peat *et al*[43] for inclusion in prognostic model protocols, and follows the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) Statement guidance[46] and the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS) checklist[49] in terms of the level of detail provided. This is to ensure provision of sufficient information to permit full assessment of risk of bias and applicability.

The aim of this study is to develop a prognostic model, the Medicines Optimisation Assessment Tool (MOAT), to identify adult patients at highest risk of preventable, moderate or severe MRPs during admission to a UK medical ward. Additional objectives are to assess the MOAT's content validity, feasibility of use, potential efficiency savings and the potential clinical risk associated with false-negative predictions. The proposed purpose of the MOAT is to permit appropriate targeting of patients by pharmacy staff in order to reduce risks, improve patient outcomes and increase efficiency of hospital clinical pharmacy services, thereby supporting delivery of national targets related to patient safety, medicines optimisation and service provision.

METHODS AND ANALYSIS

Design

This is a prognostic model development study that aims to select candidate predictors (the potential prognostic factors) and combine them into a multivariable model using logistic regression. Internal validation (bootstrapping) will be used to evaluate the performance of the

model and permit adjustment for optimism. The MOAT will be developed using a prospective cohort study involving adults admitted to the medical wards at two UK hospitals in South East England. Two study sites, Hospitals A and B, were chosen to increase generalisability of the MOAT as they have markedly different patient demographics. It is anticipated that the study will be completed by April 2018.

Eligibility criteria

Patients aged 18 years old or over will be selected by means of being consecutive admissions to the medical wards (general, emergency and elderly medicine) at the two study sites during the study period. Patients admitted to other specialties such as surgery, maternity and paediatrics will be excluded due to potential differences in the prevalence/type of MRPs in these patient groups.

Patients will be excluded if:

- their admission is for investigation only (as changes to medication will be minimal)
- they are not prescribed any medication during the admission
- their entire admission is outside of core pharmacy working hours (ie, 09:00–17:00 Monday–Friday) as these patients are unlikely to receive review by a clinical pharmacist
- their prescription is not reviewed by a clinical pharmacist during the admission (eg, a patient who is present on a study ward during core pharmacy working hours but discharged before a clinical pharmacist is able to review his/her medication).

Outcome

The outcome of interest for the prognostic modelling will be MRPs that are at least moderate in severity and preventable. Research has shown that a significant proportion of hospitalised patients will experience MRPs (eg, Blix *et al*[51] reported a rate of 81%), many of which are of limited clinical significance. We will therefore use moderate or severe MRPs as this will enable the MOAT to target patients most in need of pharmacist input in terms of risk of medication-related harm or suboptimal use. It will also ensure that the MOAT is clinically relevant and feasible to implement in terms of pharmacists' workload. Similarly, only MRPs considered as 'preventable' will be considered to ensure that the MOAT identifies patients with MRPs that are amenable to pharmacist intervention either directly or through discussion with prescribers.

The definition for MRPs that will be used is 'all circumstances involving a patient's drug treatment that actually, or potentially, interfere with the achievement of an optimal outcome'.[7 10–12]

All MRPs will be classified for descriptive reporting purposes. As there is no universally accepted classification system and perceived deficiencies with some systems, we will use the aggregated classification system recently developed by Basger *et al*.[52] This provides a comprehensive classification system based on the causes of MRPs, thereby preventing any potential confusion between MRP 'causes' and 'outcomes'.

MRP data will be identified and recorded by pharmacists at the study sites as part of their routine daily clinical assessment of patients. They will record data on all MRPs identified personally or through discussion with other healthcare professionals. The hospital incident reporting systems will also be reviewed to check for any additional significant MRPs that are not identified by pharmacy staff. Following training on the use of Basger's aggregated classification system, pharmacists will be asked to classify each MRP at the point of identification.

MRP data will be collected for all study patients from admission to discharge from hospital, or the date the study closes (2 weeks after inclusion of the final study patient), whichever occurs sooner. A study close date will be used to facilitate practicality in terms of data collection, while permitting data to be collected from admission to discharge for the majority of study patients (as the mean length of stay at the study sites is approximately 6 days).

Previous research into the detection of prescribing errors, a subset of MRPs, has shown that the observed incidence is extremely dependent on the method of detection.[53] We have chosen to use prospective identification by pharmacists for this study because (1) the purpose of the study is to develop a prognostic model for MRPs that can be identified during routine clinical practice by pharmacists; (2) it will permit the identification/inclusion of MRPs that are not routinely recorded in medical notes, such as potential prescribing or administration errors that are intercepted; and (3) it will permit the MRPs to be identified by staff personally involved in the care of the study patients, increasing clinical and practical relevance.

It is acknowledged that a limitation will be the possibility of incomplete data due to pharmacy staff being required to complete this work in addition to other routine duties. To minimise this, the principal investigator will work closely with the study sites to ensure that data collection occurs at an optimal time in terms of staffing levels and workload. Staff involved in MRP data collection will also be provided with initial training to improve the consistency and reliability of data collection. The principal investigator will review all data collection forms daily and seek clarification where needed, and provide the pharmacists with ongoing fidelity training.

We also recognise that identification of MRPs may vary depending on the knowledge, experience and skills of the pharmacists collecting data. To quantify this potential variability, a simulated 'MRP identification assessment exercise' will be developed and used in a training scenario. Each simulated MRP will be treated as having a binary outcome in terms of whether or not it is identified by each pharmacist. Fleiss' kappa will be used to calculate the level of agreement between pharmacists.

In prognostic research it is recommended that the outcome is assessed while blinded to the candidate predictors (potential prognostic factors) to prevent bias.[49 54 55]

In this study it will not be possible to blind the pharmacists collecting the outcome data to the patient's clinical information (such as age, diagnosis, laboratory results and so on) as this information will form part of their clinical assessment of the patient. Despite this, the pharmacists collecting the outcome data will not know which factors will be used as candidate predictors in order to minimise the potential for this information to influence their outcome assessment.

Following anonymisation to maintain patient confidentiality and blinding, each potential MRP will be assessed by an expert panel. Agreement will be reached by consensus on whether it is a true MRP, and then confirmed MRPs will be assessed for severity. The expert panel will comprise the principal investigator, a hospital pharmacist, a senior nurse and a consultant physician. Once MRPs have been confirmed, the panel will assess each MRP for severity. As no established grading for MRPs is available, severity will be classified using a validated visual analogue scale for medication errors,[56] used previously for this purpose by Rashed *et al*.[57] MRPs will be scored independently in terms of potential patient outcomes on a scale of 0–10, where 0 represents a case with no potential adverse effect on the patient and 10 a case that would result in death. The mean score for each MRP from the panel members will be used as an index of severity, with a score of less than 3 being considered as a minor outcome (very unlikely to have an adverse effect), a score of 3–7 will be considered as moderate (likely to cause some adverse effects or interfere with therapeutic goals, but very unlikely to result in death or lasting impairment), and a score of greater than 7 will be considered to be a severe outcome (likely to cause death or lasting impairment).

No established grading system for MRP preventability is available; therefore, we considered two possible methods: the criteria provided by Schumock and Thornton,[58] and the 'P Method'.[59] We concluded that both methods were developed for adverse drug reactions, most of which are unpreventable, whereas the majority of MRPs are inherently preventable. Neither method was therefore appropriate for the present study. Pharmacists will therefore be asked to review each MRP at the point of identification to assess whether it was preventable, expressed as a dichotomous variable of yes or no. The principal investigator will then perform a second check of all MRPs to ensure consistency. To prevent 'judgement drift' a 'case law document' will be used.[60]

Candidate predictors

Candidate predictors are the variables that predict the prognostic outcome. These can include patient demographics, clinical history, physical examination, disease characteristics, test results and treatments used.

When choosing the potential candidate predictors, various recommendations have been made:

- Predictors already reported as prognostic should be included.[61 62]
- The selection should be informed by clinical understanding (ie, expert opinion) to ensure the list is comprehensive and clinically relevant.[45 62]
- Where predictors are highly correlated (eg, weight and body mass index), only one should be selected.[62]
- Potential confounders (ie, a variable that may be associated with another predictor and the outcome) should be included to permit these to be accounted for during analysis.[63]
- Use of predictors that occur infrequently can lead to inaccurate results.[62 64]
- Candidate predictors should be:
  - available at the time when the model is intended to be used.[54]
  - clearly defined, standardised and reproducible (to enhance generalisability and applicability of study results to practice).[54]
  - have minimal measurement error (as this may dilute their prognostic value).[49]

A review of the published literature identified 59 possible predictors, but substantial variations were found between studies in terms of the strength of evidence for each predictor. This is potentially due to significant differences in study design, and the outcome measure used (namely adverse drug events, adverse drug reactions, prescribing errors and MRPs). Twenty-seven of the potential predictors were selected, based on the strength of published evidence in addition to the criteria stated above, for inclusion in a survey to obtain expert opinion from healthcare professionals and patient/public representatives. The survey was administered during April–June 2016, and a total of 247 responses were received. The results showed that the majority of the potential predictors (23 of 27) were considered 'important' or 'very important'. In addition, a significant number of additional predictors (59) were suggested.[65]

When developing a prognostic model, it is necessary to limit the number of candidate predictors used to prevent 'overfitting' or 'underfitting'. Both can lead to poor performance when the model is used in an independent data set.[55] One method to reduce the number of candidate predictors is to base the selection on the univariable association between each predictor and the outcome. This is not recommended as it results in overfitting due to selection bias,[61] and can lead to predictors being wrongly excluded from the model due to the fact that the association may only become significant after adjustment for the other predictors. It is recommended that the candidate predictors are selected a priori.[47 64] We have therefore chosen to preselect the candidate predictors for development of the MOAT (table 1) using the recommendations above. An additional consideration was the selection of predictors that form part of standard clinical data sets. This was to increase the reliability of the data and minimise the potential for missing data, and to enable the MOAT to be readily incorporated into clinical practice without the need for additional tests/measurements.

Table 1 Preselected candidate predictors for the Medicines Optimisation Assessment Tool

Variable	Details/categories	Type of measurement	Number of variables*
Demographic			
Age	Age at admission to hospital (in years)	Continuous numeric	1
Socioeconomic status	Based on the English indices of deprivation 2015 (Index of Multiple Deprivation Rank)	Continuous numeric	1
Patient-related			
Previous allergy/adverse drug reaction	Yes/No	Binary	1
Body mass index	First documented result following admission	Continuous numeric	1
Number of hospital admissions	Number of admissions to the study hospital in the previous 6 months	Continuous numeric	1
Primary diagnosis	Categorised by ICD-10 coding: ▶ Endocrine ▶ Nutritional and metabolic diseases ▶ Diseases of the circulatory system ▶ Diseases of the respiratory system ▶ Diseases of the digestive system ▶ Diseases of the genitourinary system ▶ Other (all other diagnoses combined)	Nominal categorical	6
Number of comorbidities	From hospital clinical coding data (ICD-10 codes)	Continuous numeric	1
History of dementia	From hospital clinical coding data (ICD-10 codes)	Binary	1
Medicines-related			
Number of medicines prescribed	Number of 'regular' medicines prescribed on the first full day of admission to hospital (ie, excluding 'when required' and 'once only' medicines, dietary products, non-medicated topical products (eg, emollients), wound dressings)	Continuous numeric	1
Use of 'high risk medicines'	Prescribed as a 'regular' medicine during the hospital admission: ▶ Anticoagulants/direct oral anticoagulants ▶ Therapeutic heparin ▶ Antidiabetic medication ▶ Opiates (excluding codeine, tramadol and dihydrocodeine) ▶ Aminoglycosides and glycopeptides ▶ Antibiotics (excluding aminoglycosides and glycopeptides) ▶ Theophylline and aminophylline ▶ Epilepsy medicines ▶ Antipsychotics ▶ Immunosuppressants (excluding corticosteroids) ▶ Cytotoxics ▶ Lithium ▶ Antiarrhythmics ▶ Antidepressants ▶ Other (clozapine, antiretrovirals, medicines for Parkinson's disease)	Binary (for each group)	15

Continued

Table 1 Continued

Variable	Details/categories	Type of measurement	Number of variables*
Parenteral administration route	Administration of one or more regular medicines via the parenteral route (intravenous, intramuscular, subcutaneous) during the hospital admission (excluding prophylactic low molecular weight heparins, fluid replacement therapy)	Binary	1
Laboratory results			
Renal function	Creatinine clearance calculated using the Cockcroft-Gault equation (using first documented results following admission)	Continuous numeric	1
Liver disease	Liver disease defined as ALT/ALP and/or bilirubin ≥3times normal range and/or documented liver disease Laboratory results will be the first documented results following admission Documented liver disease will be established from hospital clinical coding data (ICD-10 codes)	Binary	1
Serum albumin	First documented result following admission	Continuous numeric	1
Serum potassium	First documented result following admission	Continuous numeric	1
Serum sodium	First documented result following admission	Continuous numeric	1
White cell count	First documented result following admission	Continuous numeric	1
Platelet count	First documented result following admission	Continuous numeric	1
Total number of variables*			37

*Number of variables in relation to calculating the 'events per variable'.
ALP, alkaline phosphatase; ALT, alanine aminotransferase; ICD, International Statistical Classification of Disease and Related Problems.

All data on the candidate predictors will be collected retrospectively. Data will be obtained from the information department at the study hospitals where possible, including demographic, diagnostic and comorbidity data. Laboratory data will be extracted manually from the electronic reporting system used at both hospitals. The remaining data will be extracted manually from patient medical records. Hospital A has electronic medical and prescribing records; Hospital B has paper-based systems. Manual data extraction of laboratory data will be performed by a single data analyst at each study site, independently of the research team. Data from the patient medical records will be collected by the independent data analyst at Hospital A, but due to the use of paper-based systems at Hospital B and the need to read handwritten prescriptions, these data will be extracted by the principal investigator at Hospital B. All manually extracted data will be entered directly into an electronic database. All data will be recorded as reported, with no categorisation of continuous data.

In prognostic research it is recommended that data on candidate predictors is collected blind, in terms of knowledge of the outcome and other predictors.[49 55] This is particularly important when subjective judgement is required as it prevents the assessment being influenced, which could artificially increase the associations between the predictors and outcomes. Full blinding will not be possible for this study as the independent data analysts and the principal investigator will not be blinded to all other predictor data, and the principal investigator will not be blinded to the MRP status. It is anticipated that this will have minimal impact on the accuracy of data collection as all candidate predictors selected for this study are objective measurements that are independent of observer interpretation; therefore, subjective judgement is not required. In addition all candidate predictors will be recorded contemporaneously during the admission as part of routine care/documentation, therefore without knowledge of the MRP status. To identify any possible bias, the principal investigator will perform a double check on data entry. This will involve a double check on a randomly selected 10% sample of the 1500 study patients. Sixteen data items will be checked for each of these 150 patients, giving a total of 2400 data items. The accuracy will be calculated as the percentage of data items recorded correctly. Data entry will be refined if necessary.

Sample size calculation

Sample size is often calculated based on significance testing (power calculations), but this is not straightforward for prognostic modelling studies as there is often not a clear 'measure of effect' to power the research. An

alternative method is to calculate the sample size based on the desired precision of a sample estimate.[55] An alternative approach that is commonly used is the 'rule of thumb' of '10 events per variable' (EPV).[66] This method requires the sample size to be based on the prevalence of the outcome measure and the number of candidate predictors that will be used in model development.[49 54 55 66] Although there is debate over the optimal number of EPV, with recognition that 'the rule of 10 or more EPV is not a well-defined bright line',[64] there is agreement that models developed with less than 10 EPV need to be interpreted with caution.[55 64] The reason for the potential problem with using less than 10 EPV relates to the reliability of the model when used in a new group of patients. If a model is too closely adapted to the developmental data, it can reflect associations between the candidate predictors and outcome which are due to chance rather than true associations, known as 'overfitting' or 'optimism'.[49 55 61]

For this study the sample size has been dictated by practical considerations (funding, time available and accessibility of data at the study sites), resulting in the capacity to include 1500 (1000 from Hospital A and 500 from Hospital B), plus an additional 10% per site to allow for exclusion of patients who do not meet the eligibility criteria. We have therefore used both the precision and EPV methods to consider the adequacy of this sample size, based on an estimation of the outcome prevalence in the study population.

The outcome of interest for this study is moderate or severe, preventable MRPs in hospitalised UK patients. No estimate for the prevalence of this outcome currently exists, but Blix *et al*[61] (Norway 2004) reported that 81% of 827 hospitalised patients experienced an MRP, with approximately half of all MRPs classified as 'extremely important' or 'major' in terms of clinical significance (preventability not reported). To establish the prevalence of the outcome in the study population, we carried out pilot work involving 200 patients, and found that 39% (95% CI 32% to 45%) experienced at least one moderate or severe, preventable MRP (using the proposed visual analogue scale for medication errors and a severity score of 3 or more). Although this is consistent with Blix's work, we recognise that our estimate is based on a small sample of patients (200). In addition the MRPs were severity rated by three members of the expert panel rather than the four proposed for the main study. We have therefore chosen to use the lower CI limit as an estimate of the outcome prevalence, that is, 32%. Given an anticipated outcome prevalence of 32% and a sample of 1500 patients, we anticipate identifying 480 patients with at least one moderate or severe preventable MRP.

To consider the adequacy of the sample size using the precision method, we first established acceptable target sensitivity for the MOAT by including a question in our survey of healthcare professionals and patient/public representatives, detailed above. We proposed a target sensitivity of 90% and asked survey respondents if this was acceptable. This sensitivity was selected based on previous research to develop a 'clinical decision rule' to identify emergency department patients at risk of adverse drug events.[67] Hohl *et al*[67] used a target sensitivity of 90% as this was deemed acceptable by emergency physicians and considered feasible for implementation in terms of workload for pharmacists. A total of 237 responses were received for this question: 189 (80%) answered that 90% was an acceptable target, 21 (9%) answered no and 27 (11%) were 'unsure'. As a result we concluded that 90% is an acceptable target for sensitivity. Given the anticipated number of study outcomes and a target sensitivity of 90%, this will permit the precision of the sensitivity to be estimated with 95% CIs of ±3%, which we consider to be an acceptable level of precision in terms of clinical usefulness of the MOAT.

For the EPV method our aim would be to have at least 10 events for every variable used in model development. Given the estimate of 480 outcome events, that is to say patients with at least one moderate or severe preventable MRP, this would permit the inclusion of 48 'variables' in model development. The number of variables includes all proposed candidate predictors, interactions examined (ie, where a candidate predictor has a different association with the response depending on the value of a third variable), transformations for continuous predictors (which permits modelling of non-linear predictors) and indicator variables for categorical predictors. We do not hypothesise any interactions a priori. We will explore any potential interactions during the analytical stage to establish whether there are associations that may lead to a better understanding of the final model, but recognise the risk of overfitting caused when numerous interactions are examined, with only the strongest included in the model.[55] Similarly, we will not know whether transformations are required until we examine the linearity of the continuous predictors.

Table 1 shows the total number of variables that will be used for each candidate predictor (including the indicator variables, which are the artificial variables used to represent distinct groups within a categorical variable). We propose using 37 variables to develop the MOAT, resulting in 13 EPV, given that no interactions or transformations are required.

Data analysis plan

Descriptive analysis

Descriptive information about the sample population will be provided. This will include the distribution of the relevant characteristics of study patients, including demographic data, the distribution of candidate predictors, the ranges of continuous predictors and the amount of missing data (with possible reasons for the missingness). This is to permit an assessment of the context, case mix and setting of the study; the ranges of continuous predictors that are compatible with the MOAT; and the potential impact of any missing data. Descriptive information will also be provided on the proportion of patients with MRPs, including the severity and preventability. We

will also report the outcome frequency (namely patients with at least one MRP that is moderate or severe and preventable) across the predictor categories.

Statistical analysis

The data analysis plan has been developed prior to any analysis of the study data to prevent the potential for data-driven model development and type 1 errors.[47] The study method is informed by the recommendations of the PROGRESS partnership[42 44 47 48]; however, it is recognised that a prognostic research protocol cannot be a 'rigid blueprint'. Peat *et al*[43] state that it is neither possible nor desirable to prespecify all analysis. We therefore acknowledge that it may be necessary to modify the analysis plan or carry out additional analyses in light of new findings. Analysis will be conducted using SPSS V.24.

Missing data

Complete-case analysis can lead to selection bias[55 68] and loss of statistical power/precision.[61] We will therefore examine missing data to establish the missingness mechanism, and if we can establish that the data are 'missing at random' we will use multiple imputation (using ≥20 imputed data sets) to impute the missing values. We will then carry out complete-case analysis and compare the results with those from multiple imputation to assess for potential bias due to data that are 'missing not at random'.

Candidate predictor handling

We will first review the measurement reliability of each candidate predictor and consider excluding any that are found to be unreliable.[55] We will also consider if there are any closely related predictors that may need to be combined, or one excluded from the analysis.[55] All candidate predictors that are measured as continuous variables will be analysed on their continuous scale, that is we will not dichotomise or use categorisation as this can lead to optimistic model performance.[47 49 55 61] We have chosen to treat liver disease as a binary variable (see table 1), but this is because of the variation in liver function tests dependent on the type and stage of disease, and to be consistent with pharmacy prioritisation tools currently in use in the UK.[28] Individual continuous predictors will be examined to identify/investigate unexpected values in order to establish if they are recording errors or true outliers. We will also check for linearity of continuous predictors and use appropriate data transformation if required.[69] For categorical variables we will review the number of patients within each group and consider the need to collapse groups if there are insufficient patients to permit robust modelling.[61]

Model building

We will use multivariable logistic regression modelling to develop the MOAT. This has been chosen as the outcome is binary, and all participants will be followed up to the end of the study period. The aim is to produce a parsimonious model to increase clinical applicability while retaining reasonable predictive performance.

Backwards elimination will be used to reduce the set of candidate predictors. The Akaike Information Criterion (AIC) will be used to exclude predictors. We chose to use the AIC due to the relatively small data set (hence relatively larger p values for the predictors), as it is less likely to result in underfitting than alternative methods.[69]

Adjusting for optimism

The predictive performance of prognostic models is overestimated when assessed using the same sample data used in development (known as the apparent performance), simply because the model has been optimised for that data. To account for this 'optimism', we will assess the predictive performance of our original model using bootstrap validation. Steyerberg[69] advises that 100–200 bootstraps may be sufficient; therefore, we plan to use 200, but will increase this if it is needed to achieve a stable estimate. We can then use this to calculate a 'shrinkage factor', which will be used to adjust our original model to produce the final model/MOAT.

Creating a simplified scoring system

The final prognostic model will be used to develop a simplified scoring system or 'clinical decision rule'. These differ from prognostic models in that they indicate a specific course of action, rather than simply providing an estimate of risk, and have the advantage that they are simpler to use in clinical practice as they do not require complex calculations.

We will use the method developed by Sullivan *et al*[70] to convert the regression coefficients from the final prognostic model into a score. We will then create 'risk groups' (high, medium and low) based on the scoring system.

The development of the scoring system for the MOAT will require categorisation of the continuous predictors and the selection of appropriate cut-off points for the risk categories; we will therefore seek input from clinical experts to ensure that the grouping is clinically practical and appropriate.

Assessing model performance

We will assess the predictive performance of the final prognostic model using calibration (agreement between observed and expected predictions) and discrimination (the ability to differentiate between those who do or do not experience the outcome event). The calibration will be presented as a graph of the predicted risk of experiencing a preventable, moderate or severe MRP versus the observed risk in the study sample. Discrimination will be reported as the area under the receiver operating characteristic curve.

The predictive performance of the simplified MOAT scoring system will be reported using the classification measures: sensitivity and specificity.

Future plans

The intention of this research is to develop a prognostic model with potential to be adopted widely into clinical practice. This requires the MOAT to be clinically credible, accurate, generalisable and clinically effective in improving decision making and patient outcomes.[71]

If the initial research is successful in producing a model with good predictive performance, we plan to conduct further research to assess content validity, feasibility of use, potential efficiency savings and any potential clinical risk to patients through use of the MOAT due to false-negative predictions.

Extensive external validation, involving prospective validation in a new cohort, will also be required to further assess accuracy and generalisability before routine use of the MOAT could be recommended.[44] External validation will also provide opportunity to refine the MOAT in terms of improving the accuracy such as by updating the model[44 72] and/or simplifying the scoring system.

Following external validation we also plan to carry out implementation and impact studies to establish whether the MOAT has advantages over current practice, is compatible with (and can easily be incorporated into) practice, has the potential to change pharmacists' behaviour, has a positive impact on patient outcomes and is cost-effective. We would also investigate the possibility of incorporating the MOAT into a computerised alerting system, which would permit accurate, automated risk assessments in 'real-time', which would further support implementation.

CONCLUSION

To the best of our knowledge, the MOAT will be the first evidence-based clinical prioritisation tool to identify patients most in need of pharmacists' input (in terms of their risk of moderate or severe, preventable MRPs) while in hospital. The current research aims to develop a clinically credible and accurate prognostic model. Although further validation will be needed, we believe that the MOAT will have the potential to support pharmacists' decision making by providing objective assessments of patients' risk, thereby ultimately improving efficiency and safety.

ETHICS AND DISSEMINATION

This study has been approved by the Proportionate Review Service Sub-Committee of the NHS Research Ethics Committee Wales REC 7 (16/WA/0016) and the Health Research Authority (project ID 197298). We plan to disseminate the results via presentations at professional, academic and scientific meetings and conferences, and will submit the findings for publication in a peer-reviewed journal, adhering to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for the reporting of observational studies.[73] We will also present our findings at relevant patient/public meetings at the study sites, and work with the patient and public members of the project steering group to develop a wider public dissemination strategy.

Correction notice This paper has been amended since it was published Online First. Owing to a scripting error, some of the publisher names in the references were replaced with 'BMJ Publishing Group'. This only affected the full text version, not the PDF. We have since corrected these errors and the correct publishers have been inserted into the references.

Contributors CG is the principal investigator, and is responsible for the initial concept, study design and analysis plan. BDF and LW refined the design and analysis plan. CG applied for National Institute for Health Research (NIHR) fellowship funding, with the support and guidance of BDF and LW. CG drafted the manuscript, which was then critically reviewed by BDF and LW. All authors approved the final version.

Funding This work was supported by a Clinical Doctoral Research Fellowship award from Health Education England (HEE) and the NIHR (CDRF-2014-05-033). This article represents independent research supported by the NIHR Imperial Patient Safety Translational Research Centre. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests None declared.

Patient consent Article does not contain personal medical information about an identifiable living individual.

Ethics approval National Health Service (NHS) Research Ethics Committee.

Provenance and peer review Not commissioned; peer reviewed for ethical and funding approval prior to submission.

REFERENCES

1. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Enquiry: Mid Staffordshire NHS Foundation Trust Public Inquiry*. 2013.
2. Berwick D. A promise to learn – a commitment to act. *Improving the safety of patients in England*. 2013.
3. The Royal Pharmaceutical Society. *Keeping patients safe when they transfer between care providers – getting the medicines right*. 2012.
4. The Royal Pharmaceutical Society. *Helping patients to make the most of medicines, good practice guidance for healthcare professionals in England: Medicines Optimisation*. 2013.
5. National Institute for Health and Care Excellence. *Medicines optimisation: the safe and effective use of medicines to enable the best possible outcomes*. NICE guidelines [NG5], 2015.
6. National Institute for Health and Care Excellence. *CG138 Patient experience in adult NHS services*. 2012.
7. van den Bemt PM, Egberts TC, de Jong-van den Berg LT, *et al*. Drug-related problems in hospitalised patients. *Drug Saf* 2000;22:321–33.
8. Garfield S, Barber N, Walley P, *et al*. Quality of medication use in primary care--mapping the problem, working to a solution: a systematic review of the literature. *BMC Med* 2009;7:50.
9. Lund BC. Adverse drug events in older adults: will risk factor algorithms translate into effective clinical interventions? *Expert Rev Clin Pharmacol* 2011;4:655–7.
10. Krähenbühl-Melcher A, Schlienger R, Lampert M, *et al*. Drug-related problems in hospitals. *Drug Saf* 2007;30:379–407.
11. Committee of experts on management of safety and Quality in Health Care (SP-SQS) Expert Group on Safe Medication Practices. Glossary of terms related to patient and medication safety 2005. http://www.bvs.org.ar/pdf/seguridadpaciente.pdf. (accessed Mar 2017).
12. Pharmaceutical Care Network Europe. The PCNE Classification V 7.0 2016. http://www.pcne.org/upload/files/152_PCNE_classification_V7-0.pdf. (accessed Mar 2017).
13. The Royal Pharmaceutical Society. professional Standards for Hospital Pharmacy Services. Optimising patient outcomes from medicines. 2014.

14. Stephens M. *Hospital Pharmacy*. 2nd ed: Pharmaceutical Press, 2011.
15. The Society of Hospital Pharmacists of Australia. *Standards of Practice for Clinical Pharmacy Services*. 2016.
16. American College of clinical pharmacy. Definition of Clinical Pharmacy. https://www.accp.com/stunet/compass/definition.aspx. (accessed Mar 2017).
17. Kaboli PJ, Hoth AB, McClimon BJ, *et al*. Clinical pharmacists and inpatient medical care: a systematic review. *Arch Intern Med* 2006;166:955–64.
18. Chisholm-Burns MA, Kim Lee J, Spivey CA, *et al*. US pharmacists' effect as team members on patient care: systematic review and meta-analyses. *Med Care* 2010;48:923–33.
19. The Royal Pharmaceutical Society. *Now or Never: Shaping pharmacy for the future*. 2013.
20. NHS. *Five Year Forward View*. 2014.
21. Lord Carter of Coles. *Operational productivity and performance in English NHS acute hospitals: Unwarranted variations*. 2016.
22. Nuffield Trust. *A decade of austerity? The funding pressures facing the NHS from 2010/11 to 2021/22*. 2012.
23. The King's Fund. *Understanding NHS financial pressures - How are they affecting patient care?*. 2017.
24. East & South East England Specialist Pharmacy Services. *Prioritising pharmaceutical care delivery at ward level – Vs*. 1, 2011.
25. Health Quality and Safety Commission New Zealand. All hands on deck: prioritisation criteria. 2011. https://www.hqsc.govt.nz/assets/Medication-Safety/Med-Rec-PR/MR-Workshop-2011/MR-Workshop-All-hands-on-deck-Prioritisation-criteria-Nirasha-Parsotam.pdf. (accessed Mar 2017).
26. American Society of Health-System Pharmacists. The consensus of the pharmacy practice model summit. *Am J Health Syst Pharm* 2011;68:1148–52.
27. Dodds LJ. Optimising pharmacy input to medicines reconciliation at admission to hospital: lessons from a collaborative service evaluation of pharmacy-led medicines reconciliation services in 30 acute hospitals in England. *Eur J Hosp Pharm* 2014;21:95–101.
28. NHS England. *Transformation of seven day clinical pharmacy services in acute hospitals*. 2016.
29. Standardise MA. Upskill and scale up: how one acute trust is facing the Carter challenge. *Pharm J* 2016;297:205–7.
30. Kaufmann CP, Stämpfli D, Hersberger KE, *et al*. Determination of risk factors for drug-related problems: a multidisciplinary triangulation process. *BMJ Open* 2015;5:e006376.
31. Joint British Societies recommendations on the prevention of Cardiovascular Disease. *Risk Calculator* http://www.jbs3risk.com/pages/risk_calculator.htm (accessed Mar 2017).
32. Waterlow J. Waterlow Pressure Ulcer Prevention/Treatment Policy. 2005 http://www.judy-waterlow.co.uk/downloads/Waterlow%20Score%20Card-front.pdf (accessed Mar 2017).
33. Urbina O, Ferrández O, Grau S, *et al*. Design of a score to identify hospitalized patients at risk of drug-related problems. *Pharmacoepidemiol Drug Saf* 2014;23:923–32.
34. Onder G, Petrovic M, Tangiisuran B, *et al*. Development and validation of a score to assess risk of adverse drug reactions among in-hospital patients 65 years or older: the GerontoNet ADR risk score. *Arch Intern Med* 2010;170:1142–8.
35. Tangiisuran B, Scutt G, Stevenson J, *et al*. Development and validation of a risk model for predicting adverse drug reactions in older people during hospital stay: Brighton Adverse Drug Reactions Risk (BADRI) model. *PLoS One* 2014;9:e111254.
36. McElnay JC, McCallion CR, Al-Deagi F, *et al*. Development of a risk model for adverse drug events in the elderly. *Clin Drug Investig* 1997;13:47–55.
37. Saedder EA, Lisby M, Nielsen LP, *et al*. Detection of patients at high risk of medication errors: development and validation of an algorithm. *Basic Clin Pharmacol Toxicol* 2016;118:143–9.
38. Cottrell R, Caldwell M, Jardine G. Developing and implementing a pharmacy risk screening tool. *Hosp Pharm* 2013;71:58–60.
39. Falconer N, Nand S, Liow D, *et al*. Development of an electronic patient prioritization tool for clinical pharmacist interventions. *Am J Health Syst Pharm* 2014;71:311–20.
40. Roten I, Marty S, Beney J. Electronic screening of medical records to detect inpatients at risk of drug-related problems. *Pharm World Sci* 2010;32:103–7.
41. Hemingway H, Croft P, Perel P, *et al*. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.
42. Hemingway H, Croft P, Perel P, *et al*. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.
43. Peat G, Riley RD, Croft P, *et al*. Improving the transparency of prognosis research: the role of reporting, data sharing, registration, and protocols. *PLoS Med* 2014;11:e1001671.
44. Steyerberg EW, Moons KG, van der Windt DA, *et al*. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. *PLoS Med* 2013;10:e1001381.
45. Bouwmeester W, Zuithoff NPA, Mallett S, *et al*. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012;9:e1001221.
46. Collins GS, Reitsma JB, Altman DG, *et al*. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
47. Riley RD, Hayden JA, Steyerberg EW, *et al*. Prognosis Research Strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;10:e1001380.
48. Hingorani AD, Windt DA, Riley RD, *et al*. Prognosis research strategy (PROGRESS) 4: stratified medicine research. *BMJ* 2013;346:e5793.
49. Moons KG, de Groot JA, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;11:e1001744.
50. ClinicalTrials.gov. Development of the Medicines Optimisation Assessment Tool (MOAT) NCT02582463. https://clinicaltrials.gov/ct2/show/NCT02582463. (accessed Apr 2017).
51. Blix HS, Viktil KK, Reikvam A, *et al*. The majority of hospitalised patients have drug-related problems: results from a prospective study in general hospitals. *Eur J Clin Pharmacol* 2004;60:651–8.
52. Basger BJ, Moles RJ, Chen TF. Development of an aggregated system for classifying causes of drug-related problems. *Ann Pharmacother* 2015;49:405–18.
53. Franklin BD, Birch S, Savage I, *et al*. Methodological variability in detecting prescribing errors and consequences for the evaluation of interventions. *Pharmacoepidemiol Drug Saf* 2009;18:992–9.
54. Moons KG, Royston P, Vergouwe Y, *et al*. Prognosis and prognostic research: what, why, and how? *BMJ* 2009;338:b375.
55. Moons KG, Altman DG, Reitsma JB, *et al*. Transparent Reporting of a multivariable prediction model for individual Prognosis or diagnosis (TRIPOD): Explanation and ElaborationThe TRIPOD Statement: explanation and elaboration. *Ann Intern Med* 2015;162:W1–W73.
56. Dean BS, Barber ND. A validated, reliable method of scoring the severity of medication errors. *Am J Health Syst Pharm* 1999;56:57–62.
57. Rashed AN, Neubert A, Tomlin S, *et al*. Epidemiology and potential associated risk factors of drug-related problems in hospitalised children in the United Kingdom and Saudi Arabia. *Eur J Clin Pharmacol* 2012;68:1657–66.
58. Schumock GT, Thornton JP. Focusing on the preventability of adverse drug reactions. *Hosp Pharm* 1992;27:538–38.
59. Benkirane R, Soulaymani-Bencheikh R, Khattabi A, *et al*. Assessment of a new instrument for detecting preventable adverse drug reactions. *Drug Saf* 2015;38:383–93.
60. Franklin BD, Reynolds M, Sadler S, *et al*. The effect of the electronic transmission of prescriptions on dispensing errors and prescription enhancements made in English community pharmacies: a naturalistic stepped wedge study. *BMJ Qual Saf* 2014;23:629–38.
61. Royston P, Moons KG, Altman DG, *et al*. Prognosis and prognostic research: developing a prognostic model. *BMJ* 2009;338:b604.
62. Katz MH. Multivariable analysis: a primer for readers of medical research. *Ann Intern Med* 2003;138:644–50.
63. Hayden JA, van der Windt DA, Cartwright JL, *et al*. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280–6.
64. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. *Am J Epidemiol* 2007;165:710–8.
65. Geeson C, Franklin BD, Wei L. Identification of risk (prognostic) factors for medication related problems (MRPs) occurring during hospital admission: a survey of healthcare professionals and patient/public representatives. *Int J Pharm Pract* 2017;25:40–65.
66. Peduzzi P, Concato J, Kemper E, *et al*. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;49:1373–9.
67. Hohl CM, Yu E, Hunte GS, *et al*. Clinical decision rules to improve the detection of adverse drug events in emergency department patients. *Acad Emerg Med* 2012;19:640–9.
68. Sterne JA, White IR, Carlin JB, *et al*. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
69. Steyerberg E. Clinical prediction models: a practical approach to development, validation and updating. *Springer* 2009.
70. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med* 2004;23:1631–60.

71. Altman DG, Vergouwe Y, Royston P, *et al*. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
72. Moons KG, Altman DG, Vergouwe Y, *et al*. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ* 2009;338:b606.
73. Vandenbroucke JP, von Elm E, Altman DG, *et al*. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007;4:e297.

Open Access Miscellaneous

Correction: Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study

Geeson C, Wei L, Franklin BD. Medicines Optimisation Assessment Tool (MOAT): a prognostic model to target hospital pharmacists' input to improve patient outcomes. Protocol for an observational study. *BMJ Open* 2017;7:e017509. doi:10.1136/bmjopen-2017-017509

In reference 29 - the author's name is shown as 'Standardise MA', when it should be 'Moore A'.

Reference 41 is a duplicate of reference 42, therefore the original reference 41 is not shown. These references should read:

41. Hickson RP, Steinke DT, Skitterall C, *et al*. Evaluation of a pharmaceutical assessment screening tool to measure patient acuity and prioritise pharmaceutical care in a UK hospital. *Eur J Hosp Pharm Sci Pract* 2016. doi: 10.1136/ejhpharm-2015-000829

42. Hemingway H, Croft P, Perel P, *et al*. Prognosis research strategy (PROGRESS) 1: a framework for researching clinical outcomes. *BMJ* 2013;346:e5595.

BMJ Open 2017;**7**:e017509corr1. doi:10.1136/bmjopen-2017-017509corr1

2. Poster abstract presented at Health Services Research & Pharmacy Practice conference, April 2017

Reference: Geeson C, Franklin BD, Wei L. Identification of risk (prognostic) factors for medication related problems (MRPs) occurring during hospital admission: a survey of healthcare professionals and patient/public representatives. *International Journal of Pharmacy Practice* 2017;25:49-50. doi: 10.1111/ijpp.12368

A survey study of community pharmacy practices in providing communication support to patients with sensory impairment

C. Stewart, M. Kelly and H. Herrera

University of Portsmouth, Portsmouth, UK
malachi.kelly@port.ac.uk

The *Accessible Information Specification*, [1] published by NHS England in July 2015, acknowledges that many service users receive information from health organisations in formats which they are unable to understand. This document includes the *Accessible Information Standard* (AIS); the standard aims to make sure that people who have a disability, impairment or sensory loss are provided with information that they can easily read or understand, with support, so they can communicate effectively with health care services. From 31 July 2016, all organisations that provide NHS care were legally required to adhere to the AIS.

The aim of this study was to explore the level of awareness amongst community pharmacy staff about the AIS, and the nature of communication support they provided to patients with sensory impairments.

A questionnaire was developed based on the study objectives and literature findings. The questionnaire was sent by post, to a random sample of 100 community pharmacies, identified from the GPhC register of premises, in the geographical area of Somerset. The survey gathered anonymous data exploring the topic of interest, including: awareness of the AIS, methods used for identifying and recording sensory impairments, and strategies used for effective provision of information. Microsoft Excel ® was used to analyse responses to the closed questions using descriptive statistics to determine mean decimal values for each statement. Open comments were subject to content analysis. [2] The study was approved by institutional ethical review board.

Out of 100 questionnaires sent out, 29 were returned. Of these, 4 respondents (14%) were aware of the AIS. Observation of communication aids such as hearing aids (15 respondents (52%)), observation of behaviours (11 respondents (38%)), and being informed by the patient or carer (14 respondents (48%)) were cited most commonly as methods used to identify that a sensory impairment was present. Strategies stated most commonly for communication with those with hearing impairments were hearing loop systems (22 respondents (76%)), and for visual impairments, large print resources (22 respondents (76%)). Patient medication records were commonly used for recording of patients' sensory impairments and communication needs, 22 respondents (76%) cited this use. 9 respondents (31%) stated that they had a communications policy or guideline in place to support communication with patients with sensory impairment. The most commonly reported barriers to provision of tailored communication support to patients with sensory impairment were time constraints (14 respondents (48%)) and lack of training (10 respondents (34%)).

Community pharmacy staff awareness of the AIS is important, not only from the perspective of complying with their obligations as NHS contractors, but also in the delivery of adequate patient care. Despite the level of awareness found, many pharmacies are complying with some aspects of the standard and striving to support this patient group. There needs to be better communication between policy-makers and providers as the message regarding the AIS appears not to have got through. Further exploration of communication between policy-makers and service providers is warranted to develop ways to bridge this communications gap. This is a small scale study with a low response rate from which results cannot be considered to be representative of practices across the sector. However, it has provided valuable information to inform the direction of further research.

1. NHS England, Patients and Information. Accessible Information Specification [Internet]. Leeds: NHS England, 2015 [cited 2016 October 1]. Available from: https://www.england.nhs.uk/wp-content/uploads/2015/07/access-info-spec-fin.pdf.

2. Hsieh J, Shannon S. Three approaches to qualitative content analysis. Qualitative Health Research 2005; 15: 1277–1278.

Walk 4

Identification of risk (prognostic) factors for medication related problems (MRPs) occurring during hospital admission: a survey of healthcare professionals and patient/public representatives

C. Geeson[a,b], B. D. Franklin[b] and L. Wei[b]

[a]Luton and Dunstable University Hospital, Luton, UK
[b]UCL School of Pharmacy, London, UK
cathy.geeson@ldh.nhs.uk

There is a growing need to target hospital clinical pharmacy services, resulting in a call for the development of clinical pharmacy prioritisation tools.[1]

Our aim was to obtain opinion on the potential prognostic factors (PFs) that cause medication related problems (MRPs) during hospitalisation. This is part of a larger study to develop a prognostic model, (the Medicines Optimisation Assessment Tool/MOAT), to identify patients at highest risk of MRPs (NHS REC approval 16/WA/0016).

Potential PFs were identified from published literature, and an internet survey (comprising open and closed

questions) developed to identify: (i) the perceived importance/clinical relevance of these PFs; (ii) other potential PFs.

The survey was administered during April-June 2016. The target subjects comprised healthcare professionals and patient/public representatives. Invitations to participate were shared through various forums (e.g. Royal Pharmaceutical Society Research & Evaluation Network) and emailed to local healthcare professionals, academics, and patient/public representatives, with respondents also requested to share the survey within their organisations.

Respondents rated each PF using 5 Likert options (very important, important, 50:50, less important, not important). The median and interquartile range were calculated for each PF to establish central tendency and variability, treating responses as ordinal data.

A total of 247 responses were received (73.4% pharmacists, 13.1% doctors, 4.2% nurses, and 3.4% patient/public representatives). Due to the "infinite" target population it is not possible to determine a response rate.

Table 1 Provides a summary of the perceived importance/clinical relevance of the PFs.

Table 1 Categorisation of the perceived importance of the proposed PFs as determined by the median response

Prognostic factor	Median response*	Interquartile range
Renal function	1	0
Liver function	1	1
Age	1	1
Comorbidities	1	1
Allergies	1	1
Swallowing problems	1	1
Number of medicines prescribed	1	1
Number of potentially inappropriate medicines prescribed	1	1
Type of medicine prescribed	1.5	1
Serum sodium level	2	1
Serum potassium level	2	1
Platelet count	2	1
Serum albumin level	2	1
White blood cell count	2	2
Diagnosis/reason for admission	2	1
Type of hospital department/speciality	2	1
Readmission to hospital within 30 days	2	1
Number of hospital admissions within 6 months	2	1
Elective versus planned admission	2	1
Route of administration of medication	2	1
Dosing frequency of medication	2	1
Social deprivation	2	1
Dependent living situation	2	1
Ethnicity	3	2
Hyperlipidaemia	3	2
Number of outpatient appointments within 6 months	3	1
Gender	4	1

*Likert responses allocated ordinal numbers, 1 = very important, 2 = important, 3 = 50:50, 4 = less important, 5 = not important.

Fifty nine additional PFs were suggested, including dementia (34 participants); adherence/compliance (17); physical/sensory impairment (14); compliance aid (11); frailty (10); language barrier (9).

The study found that the majority of PFs (23/27) were considered 'important' or 'very important', with a significant number of additional PFs (59) suggested, demonstrating the multidimensional causality of MRPs.

This study enables healthcare professional and patient/public opinion to guide development of the MOAT, thereby increasing its clinical credibility.[2] It also has the potential to inform the development of other clinical pharmacy prioritisation tools. However, limitations include the use of convenience sampling, use of an "infinite" target population, and the potential impact of volunteer bias.

1. NHS England, Transformation of seven day clinical pharmacy services in acute hospitals, September 2016.

2. Bouwmeester W et al. Reporting and methods in clinical prediction research: a systematic review, *PLoS Med.* 2012;9(5):e1001221.